

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2003, 04:18:30 ; Search time 4821.31 Seconds  
(without alignments)  
15096.740 Million cell updates/sec

Title: US-09-702-134-7212\_COPY\_28000\_30500

Perfect score: 2501

Sequence: 1 atgtcaagataaaattggc.....ctatcaatggtgtctcaat 2501

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg\_hum.\*
- 31: em.htg\_inv.\*
- 32: em.htg\_other.\*
- 33: em.htg\_mus.\*
- 34: em.htg\_pln.\*
- 35: em.htg\_rod.\*
- 36: em.htg\_man.\*
- 37: em.htg\_vrt.\*
- 38: em.sy.\*
- 39: em.htgo\_hum.\*
- 40: em.htgo\_mus.\*
- 41: em.htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result	No.	Score	Match	Length	DB	ID	Description
c	1	2501	100.0	145999	2	AC096690	Oryza sat
	2	2501	100.0	244803	2	AC128647	AC128647 Oryza sat
	3	1568.2	62.7	1598	8	OSU50333	U50333 Oryza sativ
	4	602.2	24.1	1451	8	TAY14009	Y14009 Triticum ae
	5	596.8	23.9	1343	8	AB005555	AB005555 Triticum
	6	596.8	23.9	1431	8	TAY14007	Y14007 Triticum ae
	7	590.6	23.6	1446	8	TAY14008	Y14008 Triticum ae
	8	586.4	23.4	1171	8	AY014280	AY014280 Lolium pe
	9	578.4	23.1	1175	8	AY014276	AY014276 Lolium pe
	10	568.8	22.7	1167	8	AY014281	AY014281 Lolium pe
	11	499.4	20.0	2128	8	AY014277	AY014277 Lolium pe
	12	367.6	14.7	1170	8	AB077025	AB077025 Oryza sat
	13	367.6	14.7	1170	8	AY114310	AY114310 Oryza sat
	14	313.2	12.5	1513	8	PVU70531	U70531 Phaseolus v
	15	309	12.4	1511	8	AB012203	AB012203 Lactuca s
	16	299.4	12.0	1601	8	PVU70530	U70530 Phaseolus v
	17	298	11.9	1698	8	CSI250187	AJ250187 Citrus si
	18	292.4	11.7	1259	6	A41577	A41577 Sequence 3
	19	292.4	11.7	1259	6	AR138331	AR138331 Sequence
	20	292.4	11.7	1259	6	ATG0X2301	X83379 A.thaliana
	21	287.2	11.5	1857	8	PTXPTGIBB	AJ001326 Populus t
	22	282.2	11.3	1512	8	STU291453	AJ291453 Solanum t
	23	282.2	11.3	1629	8	FSY420192	AJ420192 Fagus syl
	24	280.2	11.2	1140	8	AB012856	AB012856 Nicotiana
	25	280.2	11.2	1140	8	AB016083	AB016083 Nicotiana
	26	278	11.1	1806	8	AB012204	AB012204 Lactuca s
	27	277.8	11.1	1539	8	SDU252088	AJ252088 Solanum d
	28	275.4	11.0	1728	8	AF494087	AF494087 Nicotiana
	29	271	10.8	1425	8	ATG0X2353	X83380 A.thaliana
	30	271	10.8	1490	6	A41579	A41579 Sequence 5
	31	271	10.8	1490	6	AR138332	AR138332 Sequence
	32	269.2	10.8	1508	8	MMI09112	Y09112 M.macrocarp
	33	266	10.6	1307	8	STU291454	AJ291454 Solanum t
	34	262.8	10.5	1574	8	AF494088	AF494088 Nicotiana
	35	262	10.5	1363	8	AF049898	AF049898 Lycopersi
	36	257.8	10.3	1275	8	AF049899	AF049899 Lycopersi
	37	255.2	10.2	1425	8	ATG0XV169	X83381 A.thaliana
	38	248.4	9.9	1468	8	PVU70532	U70532 Phaseolus v
	39	246.4	9.9	1435	8	AB031202	AB031202 Lactuca s
	40	245.8	9.8	465	8	AF249881	AF249881 Sorghum b
	41	243.8	9.7	1364	8	SOU33330	U33330 Spinacia ol
	42	241.2	9.6	1104	8	AB016084	AB016084 Nicotiana
	43	237	9.5	1389	8	PSIB200X	X91658 P.sativum m
	44	235.8	9.4	1392	8	AF049900	AF049900 Lycopersi
	45	235.4	9.4	1413	8	PSU70471	U70471 Pisum sativ

# ALIGNMENTS

RESULT 1  
AC096690/c  
LOCUS  
DEFINITION  
Oryza sativa chromosome 3 clone OSUNBa0059606, \*\*\* SEQUENCING IN  
PROGRESS \*\*\*, 5 ordered pieces.  
ACCESSION  
AC096690  
VERSION  
AC096690.4 GI:20503077  
KEYWORDS  
HTG; HTGS; PHASE2.  
SOURCE  
Oryza sativa.  
ORGANISM  
Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
REFERENCE  
1 (bases 1 to 145999)  
AUTHORS  
Buell,C.R., Yuan,Q., Ouyang,S., Liu,J., Tallon,L., Gansberger,K.,

Kim M., Overton II, L., Bera, J., Tsitrin, T., Krol, M., Jarrahi, B., Jin, S., Koo, H., Zismann, V., Hsiao, J., Blunt, S., Vanaken, S., Uterbach, T., Hsiao, J., Haas, B., Sub, B., Peterson, J., Quackenbush, J., White, O., Salzberg, S., and Fraser, C.  
Oryza sativa ssp. japonica cv. Nipponbare O5NBA0059G06 BAC genomic sequence

JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 145999)  
AUTHORS Buell, R.

JOURNAL Direct Submission  
TITLE Submitted (22-SEP-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA

REFERENCE 3 (bases 1 to 145999)  
AUTHORS Buell, R.

JOURNAL Direct Submission  
TITLE Submitted (08-MAY-2002) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA

COMMENT On May 8, 2002 this sequence version replaced gi:19881803.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 5 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.

\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.

\* 1 3902: contig of 3902 bp in length  
\* 3903 4002: gap of unknown length  
\* 4003 71267: contig of 67265 bp in length  
\* 71268 71367: gap of unknown length  
\* 71368 94679: contig of 23312 bp in length  
\* 94680 94779: gap of unknown length  
\* 94780 122220: contig of 27441 bp in length  
\* 122221 123230: gap of unknown length  
\* 123231 145999: contig of 23679 bp in length.

FEATURES Location/Qualifiers  
1..145999  
/organism="Oryza sativa"  
/cultivar="Nipponbare"  
/db\_xref="taxon:4530"  
/chromosome="3"  
/clone="OSNBA0059G06"  
/note="japonica cultivar-group"

BASE COUNT 40847 a 32003 c 31335 g 41412 t 402 others  
ORIGIN

Query Match 100.0%; Score 2501; DB 2; Length 145999;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCCAGATAAATTTGGCCCAAGTGAAGAACTCAGAACTACATAATAATTTTCGCA 60  
Db 115536 ATGTCCAGATAAATTTGGCCCAAGTGAAGAACTCAGAACTACATAATAATTTTCGCA 115477

QY 61 ATGCCCAAACTATTTTGGATGGTATTTGACGTGTGGCATCAGGTGCAAGTG 120  
Db 115476 ATGCCCAAACTATTTTGGATGGTATTTGACGTGTGGCATCAGGTGCAAGTG 115417

QY 121 AAAACCCATTTTTCATATGATCTGTGTAGAGGCTGTGCAAACTGATTTTTCAG 180  
Db 115416 AAAACCCATTTTTCATATGATCTGTGTAGAGGCTGTGCAAACTGATTTTTCAG 115357

QY 181 CGAGCGAAGTAAAGCTTTTCGACTAGAACGATTTTTTTTTTAAATAAATAAACAC 240  
Db 115356 CGAGCGAAGTAAAGCTTTTCGACTAGAACGATTTTTTTTTTAAATAAATAAACAC 115297

QY 241 TACAGATCCAGACCTATAGTAGGTAGCGCATGTGGCCGCGCGCTGTGGTCGGT 300  
Db 115297 TACAGATCCAGACCTATAGTAGGTAGCGCATGTGGCCGCGCGCTGTGGTCGGT 115237

QY 301 TGGCTGGGTCAAGAGACAGTGGGCTAAGCAAAATACCGGTTGTGGCACCACCTTT 360  
Db 115237 TGGCTGGGTCAAGAGACAGTGGGCTAAGCAAAATACCGGTTGTGGCACCACCTTT 114957

Db 115236 TGGCTGGGTACAAAGACAGACAGTGGGCTAAGCAAAATACCGGTTGTGGCACCACCTTT 115177

QY 361 CACATTCACCTCCCTCTTCTGATATCTTCTCTCATGGAAAGAGAGATAAGTAATTTTA 420  
Db 115177 CACATTCACCTCCCTCTTCTGATATCTTCTCTCATGGAAAGAGAGATAAGTAATTTTA 115117

QY 421 ATTGATCCCGGATAGAGAGAGAGAGAGATTAAAGAGGTAGCTAGGGAGAGAGCGAG 480  
Db 115116 ATTGATCCCGGATAGAGAGAGAGAGAGATTAAAGAGGTAGCTAGGGAGAGAGCGAG 115057

QY 481 TTGATCCCGTGTATCGATCGATCTGTGGCGCAGCGTGTATATAGGCGCGGAGAG 540  
Db 115056 TTGATCCCGTGTATCGATCGATCTGTGGCGCAGCGTGTATATAGGCGCGGAGAG 114997

QY 541 GAGTGAAGAGACAGCAGCTAGCTAGCCGCGCTGGTCCGATCCAGCTGCTGGGGATGAG 600  
Db 114997 GAGTGAAGAGACAGCAGCTAGCTAGCCGCGCTGGTCCGATCCAGCTGCTGGGGATGAG 114937

QY 601 TACTTAGTTAGCTCGGAGCTAGCTAGCTAATGAGATGATATCTATGCTAGTTAGTTAAAT 560  
Db 114936 TACTTAGTTAGCTCGGAGCTAGCTAATGAGATGATATCTATGCTAGTTAGTTAAAT 114877

QY 661 ACAGTTATTAGTTAGTTAGTTGCTATCTATCATATCTCCATCGTTAATTAATTGATT 720  
Db 114876 ACAGTTATTAGTTAGTTAGTTGCTATCTATCATATCTCCATCGTTAATTAATTGATT 114817

QY 721 GATAGCTAGTTATCAACAATTAATGAGATGGTGGTGCAGAGAGAGAGGTGGTGT 780  
Db 114816 GATAGCTAGTTATCAACAATTAATGAGATGGTGGTGCAGAGAGAGAGGTGGTGT 114757

QY 781 TCAGCGCGGGGTGCTGAGCGGGGAGAGATCCGTCGCGAGTTATATATGSCCGCGG 840  
Db 114756 TCAGCGCGGGGTGCTGAGCGGGGAGAGATCCGTCGCGAGTTATATATGSCCGCGG 114697

QY 841 AGAGAGACCCCGGGTCGGTGGCGGTGGAGAGCTGGAGGTGGCGCTGATCGAGTGGGG 900  
Db 114696 AGAGAGACCCCGGGTCGGTGGCGGTGGAGAGCTGGAGGTGGCGCTGATCGAGTGGGG 114637

QY 901 CGGGGGGAGAGAGGTGCTGCGTGGTCCGAGAGTGGGGAGAGCGTGCAGAGGACACGGCT 960  
Db 114636 CGGGGGGAGAGAGGTGCTGCGTGGTCCGAGAGTGGGGAGAGCGTGCAGAGGACACGGCT 114577

QY 961 TCTTCTGTGGTGTAAACACAGCATCGAGCGCGCTGCTGGAGAGAGCGCACCGGTGCA 1020  
Db 114576 TCTTCTGTGGTGTAAACACAGCATCGAGCGCGCTGCTGGAGAGAGCGCACCGGTGCA 114517

QY 1021 TGGACGCTTTCTTCAGCTGGCGTGGGGGAGAGAGCAGCGCGCGCGCGCGGGG 1080  
Db 114516 TGGACGCTTTCTTCAGCTGGCGTGGGGGAGAGAGCAGCGCGCGCGCGCGGGG 114457

QY 1081 AGAGCTCGGCTTACGCCAGCAGCTTACGCCGGCGCTCCGCTCCAGCTCCGCTGGAGG 1140  
Db 114456 AGAGCTCGGCTTACGCCAGCAGCTTACGCCGGCGCTCCGCTCCAGCTCCGCTGGAGG 114397

QY 1141 AGAGCTGTGCTTCCGCTACTCATCGCTGGAGTGAAGAGAGCGGAGAGGCGCGTGGGTG 1200  
Db 114396 AGAGCTGTGCTTCCGCTACTCATCGCTGGAGTGAAGAGAGCGGAGAGGCGCGTGGGTG 114337

QY 1201 AGTACCTGTGTCGGAAGCTGCGGGCGGAGCAGCGGGCGGGCTGGCGAGGTGTACTTCG 1260  
Db 114336 AGTACCTGTGTCGGAAGCTGCGGGCGGAGCAGCGGGCGGGCTGGCGAGGTGTACTTCG 114277

QY 1261 GCTACTGCCAGAGATAGACCGCTGCTGCTGGAGCTGATGAGGTGCTCGGGAGAGCC 1320  
Db 114276 GCTACTGCCAGAGATAGACCGCTGCTGCTGGAGCTGATGAGGTGCTCGGGAGAGCC 114217

QY 1321 TGGCATCTGTCGAGAGCGGGCCACTACTTCCGCGGATTTCTCCAGCGCAACACTCCA 1380  
Db 114216 TGGCATCTGTCGAGAGCGGGCCACTACTTCCGCGGATTTCTCCAGCGCAACACTCCA 114157

QY 1381 TCATCGCCTCAACTACTACCGCGGTGCCAGAGGCCACTTCGACACGCTGGGCGCCGCTC 1440  
Db 114156 TCATCGCCTCAACTACTACCGCGGTGCCAGAGGCCACTTCGACACGCTGGGCGCCGCTC 114097

[illegible]

RESULT 2	AC128647	244803 bp	DNA	linear	HTG 20-JUL-2002
LOCUS	Oryza sativa chromosome 3 clone OSUNB0062G19	***	33 unordered pieces.		*** SEQUENCING IN
DEFINITION	Oryza sativa chromosome 3 clone OSUNB0062G19	***	33 unordered pieces.		*** SEQUENCING IN
ACCESSION	AC128647.1	GI:21913026			
VERSION	AC128647.1	GI:21913026			
KEYWORDS	HTG; HTGS-PHASE1.				
SOURCE	Oryza sativa				
ORGANISM	Oryza sativa				
REFERENCE	Oryza sativa				
AUTHORS	Oryza sativa				
TITLE	Oryza sativa ssp. japonica cv. Nipponbare OSUNB0062G19 BAC genomic				
JOURNAL	Oryza sativa ssp. japonica cv. Nipponbare OSUNB0062G19 BAC genomic				
REFERENCE	Oryza sativa ssp. japonica cv. Nipponbare OSUNB0062G19 BAC genomic				
AUTHORS	Oryza sativa ssp. japonica cv. Nipponbare OSUNB0062G19 BAC genomic				
TITLE	Oryza sativa ssp. japonica cv. Nipponbare OSUNB0062G19 BAC genomic				
JOURNAL	Oryza sativa ssp. japonica cv. Nipponbare OSUNB0062G19 BAC genomic				
COMMENT	Oryza sativa ssp. japonica cv. Nipponbare OSUNB0062G19 BAC genomic				









QY	1001	GGAGGAGCGGACCGGTGCATGGAGCGCCTTCTTACGCTGCGCGTGGGGAGAGACGCG	1060
Db	361	GGCGAGCGCAGCGCTGGCTGGAGCAACTTCTTACCATGTCGCCCTCCCGGAGAGCAGCG	420
QY	1061	GGCGCAGCGCGCGCGGGGAGAGTGGCTGAGTACGCGACGACGCTTTCAGGGGCGCTTCGC	1120
Db	421	CGCCCTGGCGCGCCCGGGGAGTCTCGGGCTAGCGAGCATTTACCGCGCGGTTCGC	480
QY	1121	GTCCAAGTGCCTGTGAAGAGAGACGCTGTGTTCCGCTACTCATCGGTGGAGATGAAGA	1180
Db	481	CTTCCAAGTGCCTGTGAAGAGACACCTCTCTCTTCCGGTCTCTCCCGCTC-----	528
QY	1181	GGGCGAGGAGGGGCTGGGTGACTACTCTGGTTCGGAGAGCTCGGGCGGAGCAAGGGCGCG	1240
Db	529	CGACCCCGCCCTGTCGTGCTGACTCATGTCGCGACCCCTCGCGAGGAGCAC---CGCGG	585
QY	1241	GCTGGGCGAGGTACTTCCGCTACTGCCAGAGATGAGCCGCTGTGCTGGAGCTGAT	1300
Db	586	CCTCGGGGAGGTGTACGCGCCTACTGCTGGAGATGAGCGTCTGTGCTGGAGATCAT	645
QY	1301	GGAGTGTCTGGGAGAGCCTGGGCACTGTCGGAGACCGCGCGCACACTCTTCGGCGATT	1360
Db	646	GGAGTGTCTGGGAGAGCCTGGGGTGTGGCGCTGGCC-----ACTAACCGGCGCTT	696
QY	1361	CTTCCAGCGCAAGAGATCCCATCATCGCGCTCAACTACTACCGCGGTGCCAGAGGCGAAT	1420
Db	697	CTTTCGAGGCAACAGACTCCATCATGCGCCTCACTACTACCGCGGTGCCAGCGCGCTT	756
QY	1421	CGACAGCTGGGCAACGCTGCCGCTGGAGCCCGCACTGCTGCACATCTCCACAGCA	1480
Db	757	GGAGACGCTGGGCAACGCGGCCCATTTGGACCGCGAGTGTGCTGACATCTCTCCACAGGA	816
QY	1481	CCACGTGGGCGGCTCGGAGGTGTGGGGAGGGGCGGTGGCGGCATCGCCCTCGCCG	1540
Db	817	CAACTCTGGCGGCTTCGAGTTCACAGGAGGGCGGTGGCGGTTCATTCGACCCCGCGC	876
QY	1541	CGGGCGCTGTGTCTCAACGCTCGGCGACACTTCATGGCGCTCTCCACGCCAGGTACCG	1600
Db	877	CGACGCTCTGTGTCTCAACATCGGCGACACTTCATGGCGCTCTCGAACGGGAGGTACAA	936
QY	1601	CAGCTGCTGTGACCGGGCGGTGTCAACAGCAGGGCGCTCGCCGCTGTGCTGGCTTCTT	1660
Db	937	GAGCTGCTCTCAACCGCGAGTGTGAACAGAGGGTCCCGCAGGAATGCTGTGGCTTCTT	996
QY	1661	CCTCTGCCGAGATGGACACGCTGGTGGCGCGCGCGGAGAGCTGGTCGACGACCAACCA	1720
Db	997	CCTCTGCCGAGATGGACAAAGCTGGTGGCGCGCGCGGAGCGCTGGTGGAGCGCTCCAA	1056
QY	1721	CCGAGGGTGTACCCGAGCTTCAACGTCGGGCGCTGTGAGACTTCACGACGCCACTTA	1780
Db	1057	CCGCGCGCTTACCCGAGCTTCACTGGCGGTGTGCTGTGGACTTCACGCAAGAGCACTA	1116
QY	1781	CAGGCGGACATGGCACGCTTCAGGCGCTTCGAGCTGGCTTAATCATCATGTCACCT	1840
Db	1117	CCGGCGGACATGAAGACCTCGAGGTCTTCTCTCTGTGGATCGTCCAGCAGCAGGCG	1176
QY	1841	GCAAC 1845	
Db	1177	CCAGC 1181	
RESULT	5		
LOCUS	AB005555		
DEFINITION	Triticum aestivum mRNA for wga20, complete cds.	1343 bp	mRNA linear
ACCESSION	AB005555		
VERSION	AB005555.1		
KEYWORDS	WGA20.		
SOURCE	Triticum aestivum (strain:Rht3 Dwarf, isolate:April Bearded)		
	6-week-old dwarf plants leaf tissue cDNA to mRNA, clone_lib:lamda		
	gt10 clone:wga20.		
ORGANISM	Triticum aestivum		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta Tracheophyta;		

QY 1094 CGCAGAGCTTCACGGGGCGCTTCGGCTCCAAAGCTGCGTGAAGAGAGAGCTGCTGCTT 1153  
 Db 491 CGCAGAGCTTCACGGGGCGCTTCGGCTCCAAAGCTGCGTGAAGAGAGAGCTGCTGCTT 550  
 QY 1154 CCGGTACTCATCGCTGAGATGAGAGGCGGAGAGGCGGTGGTGAAGTACCTGCTGCG 1213  
 Db 551 CCGGTACTCATCGCTGAGATGAGAGGCGGAGAGGCGGTGGTGAAGTACCTGCTGCG 598  
 QY 1214 GAAGCTCGGGGGGAGACACGGGGGGGCTGGGCGAGGTGTACGCGTACTGCGACGA 1273  
 Db 599 CACCTCGGGGAGAGACAC---CGCCCGCTCGGGAGGTGTACGCTCGGTACTGCTCGGA 655  
 QY 1274 GATGAGCGCCCTGCTCGCTGAGCTGATGAGGTGCTCGGGGAGAGCCTTGGCGATCGTGG 1333  
 Db 656 GATGAGCGCCCTGCTCGCTGAGATCATGAGGTGCTCGGGGAGAGCCTTGGCGGTGCG --- 711  
 QY 1334 AGACCGGGCGCACTACTTCGGCGGATTCCTCCAGCGCAACGACTCCATCATGCGCTCAA 1393  
 Db 712 -----GCCCGCGCCACTACCGCGCTCTCTTCGAGGGCAACGACTCCATCATGCGCTCAA 766  
 QY 1394 CTACTACCGCGCTGCCAGAGGCGCACTGACGCTGGGCGACCGCTCGGCACTGCGACCC 1453  
 Db 767 CTACTACCGCGCTGCCAGAGGCGCTTGGAGAGCCTTGGGCGACCGCTCGGCGCTGCGACCC 826  
 QY 1454 CACCTCGCTCACTTCACCAAGGACCACTGCTCGGGCGGCTTGGAGGTGTGGCGGAGGG 1513  
 Db 827 GAGCTCCCTGACCATCTCCACAGGACCACTGCTCGGGCGGCTTGGAGGTGTGGCGGAGGG 886  
 QY 1514 GCGGTGGCGGCGCATCCGCTCCGCGCGGCGCTGCTGCTCAAGCTGCGGCGACACTT 1573  
 Db 887 CCGGTGGCGGCTCCATCCAGCGCGCGCGCTTGGGTGCTCAACATTCGCGGCGACACTT 946  
 QY 1574 CATGGGCGCTCCAAAGGAGTACCGAGCTGCTGACCGGGCGGTGCTCAACAGCAC 1633  
 Db 947 CATGGGCGCTTAAGGAGCGTCAAGAGCTGCTCCACAGAGCGGTGCTGAGACGCG 1006  
 QY 1634 GCGCGCTCGCGCTGCTGGCTTCTTCTTCCGCGGAGATGAGAGGTGTGGCGGCG 1693  
 Db 1007 GGTCCGCGGAGAGTGCCTGGCTTCTTCTTCCGCGGAGATGAGAGGTGTGGCGGCG 1066  
 QY 1694 GCGGAGGAGCTGCTGACGACCAACCGAGGTGTACCGGAGCTTCACTGCGGCGG 1753  
 Db 1067 GCGGAGGAGCTGCTGAGCGCGGCGACCGCGCTTACCGGAGCTTCACTGCGGCGG 1126  
 QY 1754 GCTGCTGAGCTTACGACGAGCGGCTACAGGCGCGGCTTCAAGGCTTCAAGGCTTCTC 1813  
 Db 1127 GCTGCTGAGCTTACGACGAGAGCTACCGGCGCGGCTTCAAGGCTTCTC 1186  
 QY 1814 CGACTGGCTTATCATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1851  
 Db 1187 CTCTGGATCGTCCAG 1224

RESULT 6  
 TAY14007  
 LOCUS Triticum aestivum mRNA for gibberellin 20-oxidase, clone S44B. 1431 bp mRNA linear PLN 10-FEB-1999  
 DEFINITION  
 ACCESSION Y14007  
 VERSION Y14007.1 GI:2222795  
 KEYWORDS gibberellin 20-oxidase.  
 SOURCE Triticum aestivum.  
 ORGANISM Triticum aestivum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Triticum.  
 1 (bases 1 to 1431)  
 Appleford,N.E.  
 Direct Submission  
 Submitted (24-JUN-1997) Appleford N.E., Plant Sciences, IACR-Long Ashton Research Station, Dept. of Agricultural Sciences, University of Bristol, Long Ashton, Bristol. BS18 9AF UK  
 2 (bases 1 to 1431)

AUTHORS Appleford,N.E.J.  
 TITLE Cloning and characterisation of cDNAs encoding gibberellin 20-oxidase from wheat (Triticum aestivum L.)  
 JOURNAL Unpublished  
 FEATURES Location/Qualifiers  
 1..1431  
 /organism="Triticum aestivum"  
 /strain="Rht3"  
 /cultivar="Maris Huntsman"  
 /db\_xref="taxon:4565"  
 /clone="S44B"  
 /tissue\_type="scutellum"  
 /clone\_lib="LambdazapScut"  
 /dev\_stage="germinating grain 3 days at 25 degrees C"  
 113..1198  
 /codon\_start=1  
 /evidence=experimental  
 /product="gibberellin 20-oxidase"  
 /protein\_id="CAA74330.1"  
 /db\_xref="GI:2222796"  
 /translation="MVQPVFDAVILSGRADIPISQFIWPEGESPTPDAAEEHLVPLDI  
 GMLSGPAAAAEVRLVGEACERHGFQVNVHIGDAELLADAHKVCNFFTMELPK  
 QALRHGECGYSASTGRFASKLPWKETLSFRSPDPALVDYIVATIGEDHRL  
 GEVYATCSMSRLSLEIMEVLGSLVGRHAFRFEFGNDSIMRLNITPPCQPLET  
 LGTGHDPDTSLTILHQNVLGVLQVHTGRWRSIRPRADAFVNIQDITFMALSNRYK  
 SCLHRAVNSRVPKSLIAFFLCPEMDKVAPPGTLVDANPRAYPDFTWRSILDTQK  
 HYRADMTLEVFSSWLVQQQPPART"  
 BASE COUNT 263 a 481 c 433 g 254 t  
 ORIGIN  
 Query Match 23.9%; Score 596.8; DB 8; Length 1431;  
 Best Local Similarity 74.0%; Pred. No. 8.3e-82;  
 Matches 827; Conservative 0; Mismatches 252; Indels 39; Gaps 4;  
 QY 749 CATGGTGTGCACAGGAGGAGGAGGTGTTCGACGCGGGGTGTGAGCGGGAGAC 808  
 Db 97 CTGTGCTCCAGTCATGTGTGCAGCGCGGTGTTCGACGCGGGGTGTGAGCGGGAGC 156  
 QY 809 GGAGATCCGCTCGCAGTTCATATGCGCGGAGGAGAGCCCGGGTGTGCGGTGGA 868  
 Db 157 GGACATCCGCTCGCAGTTCATCTGCGCGGAGGAGAGCCCGGAGAGCGCCCGA 216  
 QY 869 GGAGCTGAGAGGTGGGCGGTGATGACGACGTGGGGGCGGG-----GCGGAGAG 913  
 Db 217 GGAGCTGACAGTCGCGCATCATGACATCGCGGCGCATCTCGCGGAGACCGCGCGC 276  
 QY 914 GTGCTGCTGCTCGCGAGGTGGGAGGCGTGGAGAGCGGCTGCTCGCGGAGAGAGCT 973  
 Db 277 CGCGAGGTGACCGCGCTGCTGGGCGAGGCGCTGCGAGCGGCGAGGCTTCTTCAGGTCG 336  
 QY 974 TAACACGCGCATCGAGGCGGCTGCTGGAGAGGCGCACCGCGGCGGAGAGAGCTGCGGCTA 1033  
 Db 337 GAACACGCGTATCGAGCGGAGCTGCTGGCGGAGGAGAGCGCGCGCGGAGAGCTTCT 396  
 QY 1034 CACGCTGCGCTGCGGGGAGAGAGCGGCGGAGCGCGCGCGCGGCGGAGAGAGCTGCGGCTA 1093  
 Db 357 CACGCTGCGCTGCGGGAGAGAGCGGCGGAGCGCGCGCGCGCGGAGAGAGCTGCGGCTA 456  
 QY 1094 CGCCAGCAGCTTCACGCGGCGCTTCCGCTCCAAAGCTGCGTGAAGAGAGAGCTGCTGCTT 1153  
 Db 457 CGCCAGCAGCTTCACGCGGCGGTTCGCTCCAAAGCTGCGTGAAGAGAGAGCTTCTTCT 516  
 QY 1154 CCGGTACTCATCGCTGAGATGAAGAGGCGGAGAGGCGGTGGGTGAGTACCTGCTGCGG 1213  
 Db 517 CCGGTACTCATCGCTGAGATGAAGAGGCGGAGAGGCGGTGGGTGAGTACCTGCTGCGG 564  
 QY 1214 GAAGCTCGGGGGGAGAGAGCGGCGGCGGCTGCGGCGAGGTGTACTCGCGTACTGCGCACGA 1273  
 Db 565 CACCTCGCGGAG 621  
 QY 1274 GATGAGCGCCCTGCTGCTGAGCTGATGAGGTGCTCGGGGAGAGAGCTGCGGAGAGAGCTGCGG 1333



















```

QY 1526 CATCCGCCCTCGCCCGGGCGCTGCTGTCACAGTCGCGGACACCTTCATGCGGCTCTC 1585
|||||
Db 936 CATCAAAACCGGATGCAATGCTTTTGTGTCAATGTTGCTGATACCTTCATGGCTCTTC 995
|||||
QY 1586 CAACGCCAGTACGCGAGTCTGTCACCGCGGTCGTCACACGACGCGGCTCGCGC 1645
|||||
Db 996 AATGCGGAGATACAGAGCTGCTGTCACAGGCGAGTGGTGAACGCGAGACACAGAAA 1055
|||||
QY 1646 CTGCTGGCTTCTTCTCTCCGCGAGATGACACGGTGGTGCCTCCCGCGGAGGAGCT 1705
|||||
Db 1056 ATCTCTGCTTCTTTTGTGTCACAGAGTGGTGGTGAACGCGAGTGAAT 1115
|||||
QY 1706 GTGCGAGCACACCGAGGCTGTACCGGACTTCACGTGGCGGGGCGCTGCTGAGCTT 1765
|||||
Db 1116 AGTGGCAAACTAGGACCAAGCTGTACCGAGATTTTACATGGCCATCTCTTGTAGTT 1175
|||||
QY 1766 CACGACGCGCTACAGGCGGACATGCGACGCTTCAGGCTTCTCCGACTGCTTAA 1825
|||||
Db 1176 CACTCAAAACATTACAGAGCTGACATGAAGAAACCTTGAGGCAATTCACCACTGGCTTCA 1235
|||||
QY 1826 TC 1827
Db 1236 AC 1237

RESULT 15
AB012203
LOCUS Lactuca sativa Ls20x1 mRNA for gibberellin 20-oxidase, complete cds.
DEFINITION
ACCESSION AB012203
VERSION AB012203.1 GI:4164140
KEYWORDS Ls20x1; gibberellin 20-oxidase.
SOURCE Lactuca sativa seed germination seed cDNA to mRNA.
ORGANISM Lactuca sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae;
Lactuca.
REFERENCE
1 (sites)
Toyomasu,T., Kawaide,H., Mitsuhashi,W., Inoue,Y. and Kamiya,Y.
Phytochrome regulates gibberellin biosynthesis during germination
of photoblastic lettuce seeds
Plant Physiol. 118 (4), 1517-1523 (1998)
99063790
2 (bases 1 to 1511)
Toyomasu,T.
Direct Submission
Submitted (16-MAR-1998) Tomonobu Toyomasu, Yamagata University,
Faculty of Agriculture, Wakaba-cho 1-23, Tsuruoka, Yamagata
997-8555, Japan (E-mail:toyomasu@eds1.tr.yamagata-u.ac.jp,
Tel:81-0235-28-2861, Fax:81-0235-28-2812)
Location/Qualifiers
1..1511
/organism="Lactuca sativa"
/db_xref="taxon:4236"
/tissue_type="seed"
/dev_stage="seed germination"
1..1511
/gene="Ls20x1"
77..1228
/gene="Ls20x1"
/function="2-oxoglutarate-dependent dioxygenase"
/codon_start=1
/product="gibberellin 20-oxidase"
/protein_id="PAA37127.1"
/db_xref="GI:4164141"
/translation="MALDCMKTPSCMSLKEBELKEDQSRISYFDASVLQHEINIPQQ
FTWPNKRNPKSKSLKLEFLDLGFLSGRSSSTKEASKLVGNACQKHGFFLVNHG
VDANLSDAQTYMDLFEPLSEKQRAKAGSCGYSFTGFRSSKLPKREILISFR
FSAKRNADIVKYDFENTMGEEFVRLGKYQYCYNANSLRLSGLIMELLISLGVNRSH
FKEFPEENNSIMRLNYPRCKPELTILGTGPHCDPTSLFTLHODNVGGLEFVDNEWR
STTPNSNAFVNVNIGDIFMALSNGRYKSLHRAVYNNKTPKSLAFLCPKDKVSP

```

```

polya_site 1511
KELVDENNPRYPDFTWATLEFTQKHYRADMTNLQAFSNWVEQKTSIT"
/gene="Ls20x1"
/note="28 a nucleotides"
BASE COUNT 478 a 321 c 295 g 417 t
ORIGIN
Query Match 12.4%; Score 309; DB 8; Length 1511;
Best Local Similarity 58.1%; Pred. No. 9.3e-38;
Matches 621; Conservative 0; Mismatches 415; Indels 33; Gaps 3;
QY 775 TGGTGTTCGACGGGGGGGTCGAGCGGGGACAGGAGATCCCGTCGCGAGTTCATATGCG 834
|||||
Db 159 TCGTGTTCGATGTCATCAGTCCCTACACACAGAAACCAACATCCCAACAGTTCATATGCG 218
|||||
QY 835 CGGCGGAGGAGGCGCCGGTGGTGGGTCGGTGGAGGAGCTGGAGTGGCGGTGATGACG 894
|||||
Db 219 CCGATATGAGAAACCAACCAACCAAAATCTAAGGACCTAGAGGTTCTCTTGTATAGACT 278
|||||
QY 895 TGGGGGGGGGGGGGAGAGTGTCTGCG-----GTGGTCCGGCGAGTGGGGG 939
|||||
Db 279 TGGGAGGCTTCTTCTTGACGCTCGAGTTCAACAAAAGAGCTCCAAACTCGTTGGTA 338
|||||
QY 940 AGCGCTCGGAGAGGACAGGCTTCTTCTGTGTGTGTAAACACGGGATTCGAGGCGGCGCTC 999
|||||
Db 339 ACGCTTGCCAAAACATAGGTTTCTTCTCGTGGTAAATCATGGAGTTGATGCAACCTAA 998
|||||
QY 1000 TGGAGGAGGCGACCGGTCATGAGCGCTTCTTACGCTGCGCTGGGGGAGAGCAGC 1059
|||||
Db 399 TATCTGATGCTCAACATACATAGGAGCTGTCTTGTAGCTTCCCTTTCAGAAAGCAA 458
|||||
QY 1060 GGGCGCAGCGCGCGCGGGGAGAGCTGCGGCTACGCCAGAGCTTCACGGGCGCTTCG 1119
|||||
Db 459 GAGCTCAAAAGAAAGCGGCTGAAAGCTGTGATATGCTAGTAGCTTCACTGGAAGATCT 518
|||||
QY 1120 CACTCAGCTGCGGTGGAAGAGAGAGCTGTGCTTCCGTTACATCATCGCTGGAGATGAAG 1179
|||||
Db 519 CCTCTAACTTCTTGGAAAGAGAGAGCTATCTTTCGAGATCTC-----AGCCGAGA 569
|||||
QY 1180 AGGCGAGGAGGCGCTGGTGTGAGTACCTGTGGGAGAGCTCGGGGCGGACGACGCGGCG 1239
|||||
Db 570 AGAAGTGGCGGATATAGTAAAGGACTACITCGAGAACACAAATGGCGGAAGATTTGTC 629
|||||
QY 1240 GCGTGGCGGAGGTGACTCGCGCTACTGCCACAGAGATGAGCGCGCTGTCGCTGGAGCTGA 1299
|||||
Db 630 GACTCGGGAAGGTATCAAGAATATTGCAATGCCATAGTAGACTTCTCTTGGGATCA 689
|||||
QY 1300 TGGAGGTGCTCGGGGAGGAGCTGGGTCATCGTCGGAGACCGCGGCGGCTACTTCCGCGAT 1359
|||||
Db 690 TGGAGTTATTGGGGTTGAGCCCTTGGGTTAACGATCTC-----ACTCAAGAAAT 740
|||||
QY 1360 TCTTCCAGCGCAACGACTCCATCATGCGCTCAACTACTACCGGCGGTGCGCAGAGGCGAC 1419
|||||
Db 741 TTTTCGAAGAAATAATCTTATAATAGAGACTCAACTACTATCCACGATGCCAAACACAG 800
|||||
QY 1420 TCGACAGCTGGGCGACCGCTCGGACTGCGAGCCCGGCGGCTCGCTACCATCTCCACAGG 1479
|||||
Db 801 AGCTCACCTTAGGAACGGGCGCTCATTTGTGACCGGACATCGTTGACAAATTCACACCAAG 860
|||||
QY 1480 ACCACGTCGCGCGCTGGAGGTGGGCGGAGGCGGCTGGCGGCGGCTCGGCGCTCGCC 1539
|||||
Db 861 ATACGTTGGTGGGCTAGAGTGTCTGGACATAGTGGGCTTCCATTACACCAAT 920
|||||
QY 1540 CCGGGCGCTGCTGCTCAACGTCGCGGACACCTTTCATGCGCTTCCAAACGCCAGGTACC 1599
|||||
Db 921 CAATAGCTTTTGTGTCAATATCGTGACACATTCATGCGACTTCAATAATGAAGATACA 980
|||||
QY 1600 GCAGCTGCTGACCGGGGCTGCTCAACAGACGCGGCGCTCGCGCTCGCTGGCTTCT 1659
|||||
Db 981 AAAGTTCTTACATAGAGCTGTGGTAAACAAACAAACTCCAAAGAAAGTCACTTCTTCT 1040
|||||
QY 1660 TCCCTCTCCCGGAGATGGACAGGCTGTGGCGCGGCGGAGGAGCTGCTCCACACAC 1719
|||||

```

Db 1041 TTCATGCCCAAGAGGATAGGTGGTGGCCACCAAGGATTTGGTGAATA 1100  
 QY 1720 ACCCGAGGGTGTACCGGACTTACGTGGCGGCGCTGCTGGACTTCACGCGGCCACT 1779  
 Db 1101 ACCCAAGGGTATATCCAGATTTTACATGGGCTACTTTCTTGAGTTCACTCAGAAGCATT 1160  
 QY 1780 ACAGGGCGGACATGCGCAGCTTCAGGCTTCTCGGACTGGCTTAATCA 1828  
 Db 1161 ACAGAGCTGACATGACACCCCTCCAGCTTTTCAAACTGGGTGAACA 1209

Search completed: March 26, 2003, 14:10:32  
 Job time : 5400.98 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: March 26, 2003, 04:18:30 ; Search time 4242.99 Seconds  
(without alignments)  
15096.740 Million cell updates/sec  
Title: US-09-702-134-7212\_COPY\_10200\_12400  
Perfect score: 2201  
Sequence: 1 cgtgttcgtttgttcac.....tagctgtgtcttttgcgtga 2201  
Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.ov.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.in.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.rac.\*
- 36: em.htg.man.\*
- 37: em.htg.vrt.\*
- 38: em.sy.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
c 1	2201	100.0	145999	2	AC096690	AC096690 Oryza sat
c 2	2201	100.0	244803	2	AC128647	AC128647 Oryza sat
c 3	2034.6	92.4	207376	2	AC104433	AC104433 Oryza sat
c 4	191.2	8.7	110969	2	CNS08C70	AL731741 Oryza sat
c 5	190.6	8.7	142475	8	AP003627	AP003627 Oryza sat
c 6	190.6	8.7	153449	8	AP003238	AP003238 Oryza sat
c 7	189.8	8.6	163279	8	AC078829	AC078829 Oryza sat
c 8	188.6	8.6	113825	2	OSJN00196	AL662998 Oryza sat
c 9	186.8	8.5	164852	2	AC092075	AC092075 Oryza sat
c 10	186.4	8.5	158805	2	AC104847	AC104847 Oryza sat
c 11	186	8.5	169663	2	CNS07VFX	AL713952 Oryza sat
c 12	185.6	8.4	91552	6	AX429455	AX429455 Sequence
c 13	185.4	8.4	148441	2	AP004691	AP004691 Oryza sat
c 14	185.2	8.4	154494	2	AC120885	AC120885 Oryza sat
c 15	184.6	8.4	179486	2	AP005305	AP005305 Oryza sat
c 16	184	8.4	172468	2	AP005589	AP005589 Oryza sat
c 17	183.8	8.4	129052	2	AP003754	AP003754 Oryza sat
c 18	183.4	8.3	134518	2	OSJN00134	AL662944 Oryza sat
c 19	183.4	8.3	141428	2	AP005558	AP005558 Oryza sat
c 20	183.2	8.3	124576	2	AP004078	AP004078 Oryza sat
c 21	183.2	8.3	172159	2	AP005492	AP005492 Oryza sat
c 22	183	8.3	133595	2	AC129227	AC129227 Oryza sat
c 23	183	8.3	160174	2	AP005103	AP005103 Oryza sat
c 24	183	8.3	170000	2	OSJN00133	AL662943 Oryza sat
c 25	182.8	8.3	121799	8	OSJN00064	AL666609 Oryza sat
c 26	182.8	8.3	142081	8	AP004572	AP004572 Oryza sat
c 27	182.8	8.3	149437	2	AP005610	AP005610 Oryza sat
c 28	182.8	8.3	166679	2	AP005192	AP005192 Oryza sat
c 29	182.8	8.3	207371	2	AP005516	AP005516 Oryza sat
c 30	182.6	8.3	207371	2	AP005516	AP005516 Oryza sat
c 31	182.4	8.3	124489	2	AP003859	AP003859 Oryza sat
c 32	182.4	8.3	148285	2	AP004690	AP004690 Oryza sat
c 33	181.6	8.3	159711	2	CNS08C7X	AL731748 Oryza sat
c 34	181.6	8.3	170025	2	OSJN00031	AL666590 Oryza sat
c 35	181.2	8.2	156629	2	AC120532	AC120532 Oryza sat
c 36	181	8.2	167379	2	AP004236	AP004236 Oryza sat
c 37	180.6	8.2	67895	2	AP004160	AP004160 Oryza sat
c 38	180.6	8.2	129104	2	CNS08C92	AL732379 Oryza sat
c 39	180.4	8.2	136186	2	AP003896	AP003896 Oryza sat
c 40	180.4	8.2	151481	2	AC131343	AC131343 Oryza sat
c 41	180.4	8.2	152799	2	AC120539	AC120539 Oryza sat
c 42	179.8	8.2	85150	2	AP003454	AP003454 Oryza sat
c 43	179.8	8.2	137478	2	AP004704	AP004704 Oryza sat
c 44	179.6	8.2	130393	2	AP003565	AP003565 Oryza sat
c 45	179.6	8.2	137775	2	CNS08C80	AL731751 Oryza sat

ALIGNMENTS

RESULT 1  
AC096690/c  
LOCUS  
DEFINITION Oryza sativa chromosome 3 clone OSJNba0059G06, \*\*\* SEQUENCING IN  
PROGRESS \*\*\*, 5 ordered pieces.  
AC096690  
AC096690.4 GI:20503077  
VERSION HTG; HTGS\_PHASE2.  
KEYWORDS Oryza sativa.  
SOURCE Oryza sativa.  
ORGANISM Oryza sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
REFERENCE 1 (bases 1 to 145999)  
AUTHORS Buell,C.R., Yuan,Q., Ouyang,S., Liu,J., Tallon,L., Gansberger,K.,







	* 229779	229826:	gap of unknown length	
	* 229827	232767:	contig of 2941 bp in length	
	* 232768	232815:	gap of unknown length	
	* 232816	235017:	contig of 2202 bp in length	
	* 235018	235065:	gap of unknown length	
	* 235066	238643:	contig of 3578 bp in length	
	* 238644	238691:	gap of unknown length	
	* 238692	244803:	contig of 6112 bp in length.	
FEATURES		Location/Qualifiers		
source		1. 244803		
		/organism="Oryza sativa"		
		/cultivar="Nipponbare"		
		/db_xref="taxon:4530"		
		/chromosome="3"		
		/clone="OSJNBb0062G19"		
		/note="japonica cultivar-group"		
BASE COUNT	68502 a	54190 c	52675 g	67759 t
ORIGIN				1677 others
	Query Match	100.0%;	Score 2201;	DB 2; Length 244803;
	Best Local Similarity	100.0%;	Pred. No. 0;	
	Matches 2201;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	CCTCGTTCGTGGTGATCCCATCCAGAAATTCGTTCCAGATCCACGCCAGCAGCCCAFTTC	60	
Dd	164850	CCTCGTTCGTGGTGATCCCATCCAGAAATTCGTTCCAGATCCACGCCAGCAGCCCAFTTC	164900	
QY	61	AAGTCTCTCTCTCTCTGTATTATAGCACAGCAACAACACTAGTCTTAGCGARGCGAGACT	120	
Dd	164910	AAGTCTCTCTCTCTCTGTATTATAGCACAGCAACAACACTAGTCTTAGCGARGCGAGACT	164960	
QY	121	GCGAGACGACGAGCATGTCTCTCTCTAGCTATTAGTGTGGGAATAATAGCTCCCC	180	
Dd	164970	GCGAGACGACGAGCATGTCTCTCTCTAGCTATTAGTGTGGGAATAATAGCTCCCC	165020	
QY	181	CCACACCGCAGCAATFGCCCTCGCGGCCTCTTCCCATFACATFACATACCACATG	240	
Dd	165030	CCACACCGCAGCAATFGCCCTCGCGGCCTCTTCCCATFACATFACATACCACATG	165080	
QY	241	TCAACGAGCTGACAGCTCCCTCCGTCGTCTCGGTCTCGGTGCOCCTCGCTCCGGTTTTCC	300	
Dd	165090	TCAACGAGCTGACAGCTCCCTCCGTCGTCTCGGTGCOCCTCGCTCCGGTTTTCC	165140	
QY	301	TTGGCCGCGCACTGCGCTGCTGCCAACCCACAGCCGCTGTCGGCGCGAGAGCGCGC	360	
Dd	165150	TTGGCCGCGCACTGCGCTGCTGCCAACCCACAGCCGCTGTCGGCGCGAGAGCGCGC	165200	
QY	361	GCAGCAGCGCGCTGCTCGTGCGAATGTCGHWGGAGCGGCGCCCTCTCCTCGTCCGCTCA	420	
Dd	165210	GCAGCAGCGCGCTGCTCGTGCGAATGTCGHWGGAGCGGCGCCCTCTCCTCGTCCGCTCA	165260	
QY	421	TTATGACGGCGCCAAAGTTCCTCTCTCTACTGTCACCACTAGTTTGTCTTCC	480	
Dd	165270	TTATGACGGCGCCAAAGTTCCTCTCTCTACTGTCACCACTAGTTTGTCTTCC	165320	
QY	481	TGGAATTCCTTCCTTCATATTAAGSTTTTCTAGCAATTAATACATATATA	540	
Dd	165330	TGGAATTCCTTCCTTCATATTAAGSTTTTCTAGCAATTAATACATATATA	165380	
QY	541	TATATATGATTAATPCTAGATTCATTATCAATATATATATATATAGTAAATCTAGAAAA	600	
Dd	165390	TATATATGATTAATPCTAGATTCATTATCAATATATATATATATAGTAAATCTAGAAAA	165440	
QY	601	TCTTATAACTCTGAACACAGAGGAGTATATATATTCGCTAATGSACTAGAGTATTTGTT	660	
Dd	165450	TCTTATAACTCTGAACACAGAGGAGTATATATATTCGCTAATGSACTAGAGTATTTGTT	165500	
QY	661	TTATTATNTGGACCTTTATGAAAATACCTAGTAGAATTCACAGAACTACATCTCTTTAT	720	
Dd	165510	TTATTATNTGGACCTTTATGAAAATACCTAGTAGAATTCACAGAACTACATCTCTTTAT	165560	
QY	721	TGGTTATATATAGTAGTAATTTGCTTCTACTAGAATTTCTATTTTACTAGAGATTAATA	780	





Db	95874	TTATAGCCTCCATATTTTAAATGTATGACCGCTGACCTTTTGTTCACAGCTTTTGACCAATT	95933
QY	902	CGTCTTATTAATAAAAAATATGATCATCTATATATATTTATTTATATGACTTGCATTGCGTCATC	961
Db	95934	CGTCTTATTCGAAAAATTTATGTAATTAATCATTTATTTATTTATTTATTAATCTGATTCGTCATC	95993
QY	962	AAATATTTTTTTAAGCATGACATAACATTTTTCATATTTTGCAAAAA-----AAAAAAC	1014
Db	95994	AAATGTTTTTTTATGACATGACATAAATATTTTCATATTTGCACAAAAATTTTGAATAAAC	96053
QY	1015	GAATGGTCAAACTTATGTCGAAAAAGTCAATGGTGCATACATATTTAAATACGAGGAGGATTA	1074
Db	96054	GAATGGTCAAACTTGGTTCGAAAAAGTCAACGGCGTCATATTTAAATACGAGGAGGAGTA	96113
QY	1075	TATGGATAGCT	1086
Db	96114	TATGTGTAGCT	96125
RESULT	5		
LOCUS	AP003627	142475 bp	DNA linear
DEFINITION	Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, PAC clone:P0459B04.		PLN 21-MAR-2002
ACCESSION	AP003627		
VERSION	AP003627.3	GI:15290128	
KEYWORDS			
SOURCE	Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA, clone:P0459B04.		
ORGANISM	Oryza sativa (japonica cultivar-group)		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.		
AUTHORS	1 Sasaki,T., Matsumoto,T. and Yamamoto,K.		
TITLE	Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC clone:P0459B04		
JOURNAL	Published Only in Database (2001)		
REFERENCE	2 (bases 1 to 142475)		
AUTHORS	Sasaki,T., Matsumoto,T. and Yamamoto,K.		
TITLE	Direct Submission		
JOURNAL	Submitted (17-MAY-2001) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program; Kannonndai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan		
REFERENCE	(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/, Tel:81-298-38-7441, Fax:81-298-38-7468)		
COMMENT	On Aug 24, 2001 this sequence version replaced gi:14575596. Genes were predicted from the integrated results of the following: GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor (October 1998 version). The genomic sequence was searched against NCBI NonRedundant protein database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP. Protein homologues of the coding regions were searched against NCBI NonRedundant protein database with BLASTP2.0. ESTs represent the identified cDNA sequences using BLASTN 2.0 with the corresponding DBJ accession no. and RGP clone ID. A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative' and '-like protein'. A gene without significant homology to any protein but with EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted with a gene prediction program is classified as a 'hypothetical' protein. The orientation of the sequence is from SP6 to T7 of the PAC clone. Detailed information on overlap and assembly quality together with annotation of this entry is available at <a href="http://rgp.dna.affrc.go.jp/genomeSeq.html">http://rgp.dna.affrc.go.jp/genomeSeq.html</a> .		
FEATURES	Location/Qualifiers		
source	1. 142475 /organism="Oryza sativa (japonica cultivar-group)" /cultivar="Nipponbare" /db_xref="taxon:39947" /chromosome="1"		



```

/clone="P0459B04"
join(6392..6700,7232..7448,7556..8040)
/gene="P0459B04.1"
CDS
join(6392..6700,7232..7448,7556..8040)
/gene="P0459B04.1"
/notes="contains EST D47569(S13151)
similar to Arabidopsis thaliana protein F1216.31
unknown protein"
/codon_start=1
/protein_id="BAB63820.1"
/db_xref="GI:15290129"
/translation="MLAKEGLSPPAAYTPRRPPPPARVPASFGSGSPVAVGDLMLR
TKAGGGGFGSHSEMDLAMYDFLENGTGGGADSRASDSALSDHADNI
SYTKQGGKEKELLSVHSLFIDVIMNSDTEYRLIIDIFRSHFIARAVSDYSLNS
AVCVKQGGKIPGGDEHYIDVIMNSDTEYRLIIDIFRSHFIARAVSDYSLNS
LPVYVGTLPRLKGLHYMDAAKSLQKNSMPLPWSRLPVLQAKWHSYERIDLIH
EQDFHTASDHQALCIGHLKRLKSLQSELDTERLLMPLIKTDMKRRAKFERRRSL
SF"
15732..16972
/notes="5' LTR"
17059..173204
/gene="P0459B04.2"
17059..173204
/gene="P0459B04.2"
/notes="probably inactive due to frameshift(s) in CDS
probably inactive due to stop codon(s) in CDS
pseudogene
similar to polyprotein"
/pseudo
complement(23838..26976)
/gene="P0459B04.3"
complement(23838..26976)
/gene="P0459B04.3"
/notes="probably inactive due to stop codon(s) in CDS
pseudogene
similar to Oryza sativa chromosome 6, P0644B06.8"
/pseudo
27447..28689
/notes="3' LTR"
join(29581..30063,31083..31412,32281..32354,32618..32708,
32819..32911)
/gene="P0459B04.4"
join(29581..30063,31083..31412,32281..32354,32618..32708,
32819..32911)
/gene="P0459B04.4"
/notes="hypothetical protein
similar to Arabidopsis thaliana chromosome 3, M0C12.6"
/codon_start=1
/protein_id="BAB63821.1"
/db_xref="GI:15290130"
/translation="MATTKSFLPPPPFIALSNRPRTTLAPTNPRRRNSLAFCAS
ASGDSPTPAEGGNPLALMRRLTLPGLDYGKRSVWEGVGLEWVSGAALLALA
LAWLRGQLRAREKRYQAVFETQACGICVGTPTVRGVTGNVVRVDSLSKSIDAV
EVEDKTLIVPRNSVBNVSGILMETLIDITPKDPLTPSVGLPDCSKEGLLICDK
EMKQCGVSLDALGVTFRIGRDMEEIGVHKYLAEKVASIMEEAPLISVRVDS
LRKVEIIRAGADAGDURLKLSMLTPENTDLIKQSIITFLTKANIESISSDSSG
FTGDDATFRNKKLLKLSRL"
join(33620..33733,35242..35382)
/gene="P0459B04.5"
join(33620..33733,35242..35382)
/gene="P0459B04.5"
/notes="hypothetical protein"
/codon_start=1
/protein_id="BAB63822.1"
/db_xref="GI:15290131"
/translation="MAHSGFGSLTSPREFLAYDMGHPLNRTVDGFLKIGAKNMGV
GAYGALVSASNSHQNVKATVGGATAEFLNYLT"
join(36642..36706,36726..36840,36947..37117)
/gene="P0459B04.6"
join(36642..36706,36726..36840,36947..37117)
/gene="P0459B04.6"
/notes="contains ESTs AU031375(E61447),AU029988(E50438)
unknown protein"

```

---

```

/codon_start=1
/protein_id="BAB63823.1"
/db_xref="GI:15290132"
/translation="MSCCGNCGSGSSCOGNGCGSDRCNSNLTIDICRCKYSEVE
PTTTTFADATNKGYYWCCFRIRROGGEROLRQHLQVRHQLRLLLQLLEETYL
HPILHINELPVPLMCTSYLACQMYLNLLAS"
complement(join(39014..39084,40410..40797,41574..42182,
42726..43001,43610..43798,44035..44145,44725..44905,
45677..45853,46704..46958,47100..47606,47693..47743,
47836..47915,48207..48236))
/gene="P0459B04.7"
/notes="hypothetical protein
similar to Arabidopsis thaliana chromosome 3, T2E22.10"
/codon_start=1
/protein_id="BAB63824.1"
/db_xref="GI:15290133"
/translation="MAVOVESKRDYIAMPSTIDMAYNVLIDHALAERDRSPAVVPKC
VSLIKTYLRYTPRVOTLROIDLFCATIAKCEPIGTORSSASPSHSSVAAPITNFA
SPSLVSLNYSVARHIPKLSFORIGHSEVAPTSTKQSLPSLSFENKSLVOLTPE
AITNMDLVESKESHAPDLISSATERADGGEPADDTKIISFDILNWHVYGERQASA
STKESNDFADLQDFHTQFLEVGAAALVGDMEAKINDQWQYSFIDFDPDIDLQIP
STSTASTFASQSHLKAITASKRMKSGNOVCEQQLKINPAEISYIAVACSETSN
SNQFSAPRLTQSQPSADYAFSVLIKVIDMTKAFDILINLGHALLHELMIVENA
PLIEKSETNHSYMNNEYSDPORATESEORMEEROETVWASALSCLFVFCDG
GKIIRSLGGLDIRVITKLLSEISVESHAKVHSLICMLAMLYOYSDGAPGALDT
HFDPQLDRVGGVDYICLEYSRANSREKRDLFVLYDYVHLQINERFLAGLSITYT
DDAQLASLLACADAPAFYISVRKHVEGVDMLRAISSALSSQTESYEQNLVILDKV
IRKLDVTSTFRIDETPAYMIQVTSYKFSIRSDGHEDADVALRCLRWATYLSHLL
NSQISYRHHGYIMLVELLSEISETDGSINWKIQLODEIVAGSDLSSEVSLP
VCLLGLLKSXNFIRWGLYVLDKFLMRCLLDDNDMOGHVTHSKHRLDKAFAY
IDMNSALLVYONNETDHIINLKMCDMLFSQLCLRUPSSNMVHMGGLSGLGOLFCT
TKNIESHLETASHONVGNKFCRSETHLODISVNOCTACTTLLSETSMALLLGLAIA
PMQLVARVPTSLFFWPLMOLGAASDDIALGIAVGTSGESFVTKLITGSKRCGYLN
EV"
53255..54595
/gene="P0459B04.8"
53255..54595
/gene="P0459B04.8"
/codon_start=1
/product="putative receptor-like kinase Xa21-binding
protein 3"
/protein_id="BAB63825.1"
/db_xref="GI:15290134"
/translation="WGHGLSCSDTDEYDLFRAOLGDIALSALLAADPALARRATV
YDRFTALHAAANGRLQVLSMLLDRDGDVLSKKQTFPLVAMRGTECVYRLRG
GANVLTDFSPRARTCLHAAVYHAECLQAILGAQAQGVAAASWGFARFVVRDER
GATPLHARHARASCVRLLDKGAIVSAPTAVYGPSTALHAAAGNMECIRBL
AWGADRLQDRASRIAYAVMRGRACAAALNPAAP IVPSPKFTGELADAKA
LLEAALMEANREKRILHGSDINIKGDEEESDEEACNICPEQACMEVKEGCH
QMACACTIACCHSNPKTLLHPACPCRTTISRLVATTNSKTNRSRSGRS
SSFGKGLSSNMGSSFRIGGSRVLVDGSSGELAKPDHDFSSVAAAALCOT"
complement(join(56259..56384,57157..57234,57351..57401,
57530..57550,58700..58869,59089..59518))
/gene="P0459B04.9"
complement(join(56259..56384,57157..57234,57351..57401,
57530..57550,58700..58869,59089..59518))
/gene="P0459B04.9"
/notes="hypothetical protein"
/codon_start=1
/protein_id="BAB63826.1"
/db_xref="GI:15290135"
/translation="MPLFSAPALHRRRLRHAAAFALVLLAVALLFLLIIVHPSLGT
PSPSPSGHRLPTLDILVHGAKEKAGVLDGTPPGYHWFPGTGDGSKNWLHLBGG
SWCRNRSYCDHRKRTSLGSSATMETKVEFGLSDDKAQONPTIETPHRTSFHFLIC
STVLRHSYIGATGATFFRGRIWEAVMAELLPKGLARAKQGHFWKDLRFRFGSS
DMEPGCVFPQEVAKGITTPFILNPAYDVMTQVAEAVGDFPDRREVKEDICEYCN
PTCFNVLYEQPYQEG"
join(65394..65405,65622..65777,66253..66420,66764..67036)

```



```

/ gene="P0459B04.10"
join(65394..63405,65622..65777,66253..66420,66764..67036)

Query Match
Best Local Similarity 8.7%; Score 190.6; DB 8; Length 142475;
Matches 222; Conservative 0; Mismatches 24; Indels 7; Gaps 1;

QY 838 CTTATTACTACCTCCATATTTTAAATGATGACGCGGTCACTTTTGTCCACAGTTTGAC 897
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 44370 CTTAATPACTACCTCATATTTTAAATGATGACGCGGTGACTTTTGTCAACATTGAC 44429
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 898 CATTCGCTCTATTAAATAAATATGATCTATTTATTTATTTTATTTATGACTTGATTCGT 957
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 44430 CATTCGCTCTATTAAATAAATATGATCTATTTATTTATTTTATTTATGACTTGATTCGT 957
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 958 CATCAAAATATTTTAAAGCATGACATAAACATTTTCATATTGCAAAA-----AAAA 1010
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 44490 CATCAAAATGCTTTAAAGCATGACATAAATATTTTCATATTGCACAAAAATTTTGAATA 44549
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 1011 AAACGAATGGTCAATGTTAGTGAAGAGTCAATGGTGCATCATTAATAACGAGGG 1070
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 44550 AAACGAATGGTCAACGTTGTCGAAAAGTCAACGGCGCTACATTAATAACGAGGG 44609
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 1071 ATTATATGGGATA 1083
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 44610 AGTATATGATAA 44622
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 6
AP003298
LOCUS 153449 bp DNA linear PLN 17-APR-2002
DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,
PAC clone:P0698H10.
AP003298
AP003298.3 GI:20160921
Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
clone:P0698H10
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzoideae; Oryza.
1
Sasaki, T., Matsumoto, T. and Yamamoto, K.
Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:P0698H10
Published Only in Database (2001)
2 (bases 1 to 153449)
Sasaki, T., Matsumoto, T. and Yamamoto, K.
Direct Submission
Submitted (19-FEB-2001) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program; Kaunondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
On Apr 16, 2002 this sequence version replaced gi:16416398.
Genes were predicted from the integrated results of the following:
GENSCANL.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor
(October 1998 version). The genomic sequence was searched against
NCBI NonRedundant Protein database, nr
(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at
RGP. Protein homologies of the coding regions were searched against
NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent
the identified cDNA sequences using BLASTN 2.0 with the
corresponding DBJ accession no. and RGP clone ID.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-' and '-like protein'. A gene without
significant homology to any protein but with EST homology (covering
almost the entire length of partial sequence) is classified as an
'unknown' protein. A gene predicted with a gene prediction program
is classified as a 'hypothetical' protein.
The orientation of the sequence is from SP6 to T7 of the PAC clone.

```

This sequence of P0698H10 clone has an overlap with P0518C01 (DBJ: AP003277) clone at the position 1 to 80,700 of 5' end and with P0459B04 (DBJ: AP003627) at the position 103,851 to 153,449 of 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://rgp.dna.affrc.go.jp/GenomeSeq.html>.

FEATURES	source
1..153449	Location/Qualifiers
/organism="Oryza sativa (japonica cultivar-group)"	
/cultivar="Nipponbare"	
/db_xref="taxon:39947"	
/chromosome="1"	
/clone="P0698H10"	
complement(join(4643..4651,4721..4846,4888..5021,8292..8312,9406..9448))	
/gene="P0698H10.1"	
complement(join(4643..4651,4721..4846,4888..5021,8292..8312,9406..9448))	
/gene="P0698H10.1"	
/note="hypothetical protein"	
/codon_start=1	
/protein_id="BAB89858.1"	
/db_xref="GI:20160922"	
/translation="MLAKOCNAMOSDOLFWMKDEVTRALGRILSGVGRVNDPLAWLAMYFSANNPLLAIPFFGKLKQCHWADLLKRGINSHPYPPYAOBLETANHLILD CVFARQW"	
complement(join(11376..11773,12758..12809))	
/gene="P0698H10.2"	
complement(join(11376..11773,12758..12809))	
/gene="P0698H10.2"	
/note="hypothetical protein"	
/codon_start=1	
/protein_id="BAB89859.1"	
/db_xref="GI:20160923"	
/translation="MEEDEDEEEKKKERKRVELISNEDRWKLSKSGRSGRSCSSSNYYHSDFNSSKNATVLSRSYASVTSRHTATWAGAGGGAGSSSSSQ HQHQOQQQSNRSORLSKCVKVAEKEHAEFYIVRCVSNLMYCRDY"	
complement(join(16198..16340,16376..16523,17419..17541,19928..19954,20609..20764))	
/gene="P0698H10.3"	
complement(join(16198..16340,16376..16523,17419..17541,19928..19954,20609..20764))	
/gene="P0698H10.3"	
/note="hypothetical protein"	
/codon_start=1	
/protein_id="BAB89860.1"	
/db_xref="GI:20160924"	
/translation="MHAGTDRFKSCDVLAAANSQDTLCTVLAWRRAHAYVVLSSV HLLIORTYATKELKNKLASYLIDIRFFEFFEFFEFFELVSLTSCVYDIPGLPQ RWMHAKASRLDAVGGINQOSRILRASIRYDMYADRSMLPASCVRYYIYCYAV EHC SIQVRVGVVVCISHIGVISISINENHNSPNYK"	
complement(join(26367..26488,26778..26908,27759..28784,28412..28548,29329..29416))	
/gene="P0698H10.4"	
complement(join(26367..26488,26778..26908,27759..28784,28412..28548,29329..29416))	
/gene="P0698H10.4"	
/note="hypothetical protein"	
/codon_start=1	
/protein_id="BAB89861.1"	
/db_xref="GI:20160925"	
/translation="MQGLMSVFVIIIGNYALDWITDVIYFMIGVGMHLRCOTVKL IIAATFLYFVMAFGFKQLIEYNIAFCPTVFGRKSTRETMTKNIKALGFGYI WFERLSFLRSCCCFFFLVVVVVVVIAWMTQTQAVETNDIQAGTFWLDGPGAS RAEPRPAGATGICSLIARAYRCPLRMHACTTYCA"	
complement(join(32396..33552,33673..33814,34064..34197,35201..35214,36092..36171))	
/gene="P0698H10.5"	
complement(join(32396..33552,33673..33814,34064..34197,35201..35214,36092..36171))	
/gene="P0698H10.5"	
/note="hypothetical protein"	
/codon_start=1	

```

/protein_id="BAB89862.1"
/db_xref="GI:20160926"
/translation="MDRIIGSSNTTMRARSTTDNRYIRMSDRSGISPCSDVILV
SVGQDAEVDALMHRDLFRVLTIDMNLSSALQPSRVYIMPLTLGRVADGFR
LTLIAALRKLIDHVSQPRHDTAVIYVFDRLPPTPGTVSACIAGLRCRWA
SSIPACLRRLRVISPMALGHTTIVHGRPLRHPAGLRRRLVFPARLAYACAA
GHLRHRFTSYAADWCLRLHGWPIIMPPLIGVDFTAGLPPSPSGASVATGILLCRR
CWASTSPACLRRCRVVPPPTFLAYAAAVGRRLRHRPASVAADWCPRLGWLTP
PMGVHVPADRLAPAAADWCFRLHGWPIIMPPLIDVYDFTAGLPPSPPTGLVATVGN
PMSHVASSTSPGTPSPPPVGVSAITSLCLRCWASTSPCLACLRRLVPPFRLA
YIAAAGRLRLHRRPASVAADWCPRLIGWLVTTPFMGVHVPADRLAPAAADWCFRLH
"
gene join(40057..40430,40631..40783,41168..41861)
/genes="P0698H10.6"
CDS join(40057..40430,40631..40783,41168..41861)
/genes="P0698H10.6"
/notes="hypothetical protein"
/codon_start=1
/protein_id="BAB89863.1"
/db_xref="GI:20160927"
/translation="MORSREYVRDGDVYGEELLREILDETAAVHSNSNSNSNS
KEAEEDREYFAAADAEOLOVEAPCGRRRRESMVNKLSTVYSGTISDIESALP
TAGDQLLADHNPASSCPVFSPEYLSKTMENKYLAMKSNNGGLADGYK
WRIGKSKNSPNRYSYRCTNRCNAKQVERAVDEPDLTVTYEGLHYTSHF
LHSTSSSTTQOQOPQMNTNCKKPKLHLPLLDHDDPPPPPPPEMTTMMIM
QSFIOQQDDDLLOPAADHLVQAPDPDCYNINGSSSGLMMSLDLDDQAAAG
GLEDVVLVLRPPPPPCNNNNYYPATTCISDNEYSSASAPSSSVSYSWTTP
MSPCIDMTLSNIF"
gene join(42534..42746,42822..43280,43389..43616,44080..44259,
44355..44655,44769..44834,45256..45308,45405..45491)
/genes="P0698H10.7"
CDS join(42534..42746,42822..43280,43389..43616,44080..44259,
44355..44655,44769..44834,45256..45308,45405..45491)
/genes="P0698H10.7"
/codon_start=1
/product="putative cell cycle switch protein"
/protein_id="BAB89864.1"
/db_xref="GI:20160928"
/translation="MLMGRPAWQREXNGYSGGPTVGRQVLVLEKVDLPTPKVTVA
TSPLFLLLVVVVVGASLDVPAAPLPRNVPAMAGGLRLDVAVASPARILLD
VPTPSPSTLSDYDFEIPCRSSSLHNFALDRASPSTTDDAPYSLRLAFLGP
DSFSPAPSPNNLFRFKTHPSKSPFPAASAAATAGHDCTAGSASSTPRPKRV
PKTHVLDPASLDLYNLVDWSONTLAVGLNRCVYLWSASNCVKLCLDLGPRD
SCAVHWREGSYLAIGTSLGVDQINDSRCKRIKRNMGHGHTGTGLVAMSRLISGS
RDNILQHDIRVPSDYTSKFSGRSHSNVCASSDFQGVQGLAWHDDRELASGND
NOLLVWQSQOPIRLTHTAAVKAIAWSPHQGLLAGSGGTADRCIRWNTVNGNM
LNSVDPGSQAOTCEHSWPKENHGOVEPIVVRGYSNWTAAALPCNVTIIVTGRGD
ETLRFWNIEPSMKTVGTYC"
gene join(47888..47948,48328..48393,48479..48576,48699..48736,
48808..48899,48960..49081,49175..49252,49349..49414,
49503..49658)
/genes="P0698H10.8"
CDS join(47888..47948,48328..48393,48479..48576,48699..48736,
48808..48899,48960..49081,49175..49252,49349..49414,
49503..49658)
/genes="P0698H10.8"
/notes="contains ESTs D15121(C0112),C97914(C0112)
unknown protein"
/codon_start=1
/protein_id="BAB89865.1"
/db_xref="GI:20160929"
/translation="MECSNYSKEQRUNGSLVFLDPLTLPVTSGGADVNRTOAY
MIBKLGVSSISLECLVLLYKQXTMAGLAVGYOFDYDDFHSYHGRVLAETKRPD
PVLNRTLLSLPRKYVYENCDFQVFTNGRIHASRLKRLGIEDCFERYVCFELNPS
PILCKNIDAMQALKVASINPKTSILFDLSARNIOAKLIGMTVLVGTSEIKGAD
HALESNNKEALPELDEAVKDEDVKSKYGIETVIA"
gene join(52973..53089,53413..53571,53703..53789,53936..54136,
54208..54381,55678..55731,55862..56002,57568..57924)
/genes="P0698H10.9"
CDS join(52973..53089,53413..53571,53703..53789,53936..54136,
54208..54381,55678..55731,55862..56002,57568..57924)
/genes="P0698H10.9"
/notes="hypothetical protein"

```

```

/codon_start=1
/protein_id="BAB89866.1"
/db_xref="GI:20160930"
/translation="MVLPRQWRRRLSLRTLAASYYVLOGSDVPAGPYVYDVALLL
LVALLFSLPSSEAEAAATNLRREQVQSLRLINKPKPLATIQSPDGIIDCV
HISKOPADPLLLKNTIOPSGMYGAARPTQTNQNGKCPDNTVPIRRTKEDDM
RATSVATGKTKTHGSHPRLAGVTGDHGYGVASATGYGTATINLQPTIATS
GDFSLAOLWISAGSYONKDLNTEAGWOVYPALYDEKTRFFIYVTRDAVDQTCYNL
ACSGFIQNTYIGVSTSPVSIYGGPOYEDYLVKDLAGNWLVOVGKTYGVWESSI
FHLQGVADTVWEGGVNSFRSTTFWGSCHFPFKGFGKATSKAIQVVDSSNKLKSP
NGVSLAPLPNYINMVTSSSTTSWGTIYIYGGGCP"

```

```

Query Match      8.7%; Score 190.6; DB 8; Length 153449;
Best Local Similarity 87.7%; Pred. No. 11e-22;
Matches 222; Conservative 0; Mismatches 24; Indels 7; Gaps 1;
QY 838 CTTATTACTACCTCCATATTTAAATGTATGACGCCGTCACCTTTTGTCCACAGTTTGCAC 897
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 148220 CTTAATACTACCTTCATATTTAAATGTATGACGCCGTCACCTTTTGTCTAACATTTCAC 148279
QY 898 CATTGCTCTTATTAATAAATAATATGATCTATCTATTATTATTATTATGACTGATTCGCT 957
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 148280 CATTGCTTATTAACAAAATAATATGATCTATCTATTATTATTATTGTCGACTGATTTGT 148339
QY 958 CATCAATATTTTATTAAGCATGACATAAACAATTTTCATATTTCGCAAAA-----AAAA 1010
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 148340 CATCAATGTTCTTTAAGCATGACATAAATAATTTTCATATTTCGACAAAAATTTGAATA 148399
QY 1011 AAACAAATGTCAAATGTATGTCGAAAAGTCAATGTCATACATTAATAATACGGAGGG 1070
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 148400 AAACAAATGTCGCAACGTTGGTCGAAAAAGTCAACGGCGTCATACATTAATAATACGGAGG 148459
QY 1071 ATTATATGGGATA 1083
||||| ||||||| ||
Db 148460 AGTATATGATRAAA 148472

```

```

RESULT 7
AC078829/10 163279 bp DNA linear PLN 03-JUL-2002
LOCUS Oriza sativa chromosome 10 BAC OSJNBa0026012 genomic sequence,
DEFINITION complete sequence.
ACCESSION AC078829
VERSION AC078829.10 GI:21686922
KEYWORDS HTG.
SOURCE Oriza sativa (japonica cultivar-group).
ORGANISM Oriza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
AUTHORS Buell,C.R., Yuan,Q., Moffat,K.S., Hill,J.N., Burr,P.C., Hsiao,J.,
Zismann,V., Pal,G., Bowman,C.L., Fujii,C.Y., Vanaken,S.E.,
Bowman,C.L., Craven,B., Utterback,T.R., Khalak,H., Feldblyum,T.V.,
Quackenbush,J., White,O., Salzberg,S.L. and Fraser,C.M.
Oriza sativa chromosome 10 BAC OSJNBa0026012 genomic sequence
Unpublished
REFERENCE
AUTHORS Buell,R.
TITLE Direct Submission
JOURNAL Submitted (04-AUG-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
REFERENCE
AUTHORS Buell,R.
TITLE Direct Submission
JOURNAL Submitted (27-FEB-2001) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, rbuell@igr.org
REFERENCE
AUTHORS Buell,R.
TITLE Direct Submission
JOURNAL Submitted (13-MAR-2001) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, rbuell@igr.org
REFERENCE
AUTHORS Buell,R.
TITLE Direct Submission
JOURNAL Submitted (13-MAR-2001) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, rbuell@igr.org

```

## AUTHORS

Buell, R.

## TITLE

Direct Submission

## JOURNAL

Submitted (29-AUG-2001) The Institute for Genomic Research, 9712  
Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org  
6 (bases 1 to 163279)

## AUTHORS

Buell, R.

## TITLE

Direct Submission

## JOURNAL

Submitted (03-JUL-2002) The Institute for Genomic Research, 9712  
Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org  
On Jul 3, 2002 this sequence version replaced gi:13129494.  
Address all correspondence to:rice@tigr.org

## COMMENT

BAC clone OSJNBA0026012 is from Oryza sativa chromosome 10.  
The orientation of the sequence is from SP6 to P7 end of the BAC  
clone.

Genes were identified by a combination of several methods: Gene  
prediction programs including Genscan and Genscan+ (Chris Burge,  
<http://CCR-081.mit.edu/GENSCAN.html>), GeneMarkHM (Mark Borodovsky,  
<http://genemark.biology.gatech.edu/GeneMark/>), Fgenesh  
(<http://www.softberry.com/>), and GeneSplicer (Mihaila Pertea and  
Steven Salzberg, [contact.mpertea@tigr.org](http://contact.mpertea@tigr.org)), searches of the  
complete sequence against a peptide database and the plant EST  
database at tigr (<http://www.tigr.org/tdb/tgi.shtml>). Annotated  
genes are named to indicate the level of evidence for their  
annotation. Genes with similarity to other proteins are named after  
the database hits. Genes without significant peptide similarity but  
with EST similarity are named as unknown proteins. Genes without  
protein or EST similarity, that are predicted by more than two gene  
prediction programs over most of their length are annotated as  
hypothetical proteins. Genes encoding tRNAs are predicted by  
tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>).  
Simple repeats are identified by repeatmasker (Arian Smit,  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>).

This BAC overlaps with rice BAC OSJNBB0044B19 (AC078893) and  
OSJNBA0059P24 (AC037425).

## FEATURES

## source

Location/Qualifiers

1..163279

/organism="Oryza sativa (japonica cultivar-group)"

/cultivar="Nipponbare"

/db\_xref="taxon:39947"

/chromosome="10"

/map="E3243"

/clone="OSJNBA0026012"

/note="japonica cultivar-group"

complement(1575..1683)

/rpt\_family="At\_rich"

1643..1669

/rpt\_family="AT\_rich"

complement(2060..2100)

/rpt\_family="AT\_rich"

complement(2700..2723)

/rpt\_family="AT\_rich"

complement(5159..5181)

/rpt\_family="AT\_rich"

complement(6896..6927)

/rpt\_family="AT\_rich"

complement(11673..11753)

/rpt\_family="AT\_rich"

complement(13043..13081)

/rpt\_family="AT\_rich"

complement(16661..16686)

/rpt\_family="AT\_rich"

complement(18032..18065)

/rpt\_family="AT\_rich"

complement(26795..29035)

/gene="OSJNBA0026012.16"

/note="similar to NAM GB:CAA63101 (Petunia x  
hybrida)"

GI:4803672 (Homo sapiens)"

complement(join<26795..26803,27038..27076,27396..27443,

27721..27783,27886..27963,28907..>29035))

/gene="OSJNBA0026012.16"

complement(join(26795..26803,27038..27076,27396..27443,

27721..27783,27886..27963,28907..29035))  
/gene="OSJNBA0026012.16"  
/codon\_start=1  
/product="putative nuclear transport receptor"  
/protein\_id="AAK13152.1"  
/db\_xref="GI:13129498"  
/translation="MEQATAAVKALAAALVHHDDATTAADRWLQQFQHTLDAQV  
ADSLHDESSNNETQIFCSQTKSRKQVDFEELPSEAFRLQSLYALKKFSKGPQK  
VNTQICATAMALAVHFFG"  
29166..29196  
/rpt\_family="(TGGGGG)n"  
complement(29966..30049)  
/rpt\_family="AT\_rich"  
/rpt\_family="AT\_rich"  
33907..33945  
/rpt\_family="(CGG)n"  
complement(34290..34356)  
/rpt\_family="(TA)n"  
34452..34525  
/gene="OSJNBA0026012.LRNA-Asn-1"  
34452..34525  
/gene="OSJNBA0026012.LRNA-Asn-1"  
/product="tRNA-Asn"  
complement(35292..35312)  
/rpt\_family="AT\_rich"  
complement(36223..37996)  
/gene="OSJNBA0026012.6"  
/note="similar to NAM GB:CAA63101 GI:1279640 (Petunia x  
hybrida)"  
complement(join<36223..36759,36892..37157,37825..>37996))  
/gene="OSJNBA0026012.6"  
complement(join<36223..36759,36892..37157,37825..37996))  
/gene="OSJNBA0026012.6"  
/codon\_start=1  
/product="putative NAM (no apical meristem) protein"  
/protein\_id="AAK13151.1"  
/db\_xref="GI:13129497"  
/translation="MAGLREMESTLPPGFRCPDSEELICFYLKNKVAHVASGLTV  
DYDLHAPLPELPAKLAETAEWIFFDFDRKIATGSRNTRATGIVKATGDRIVH  
EGTTRAVYGMKRLVLYLGRAPNGKQTWVHEFRLETPNSQPKEDWLCRVDFKKP  
STIEAGGGSGSLDFPGATGSDTPSPPTMAPLLIGSPDPTVYDFRHSAAVPP  
LAVLMQGGDDMGISGVHCSNNDSGSSALLNLTMLOYSFLEHPTGDDMAVGAHF  
GTCGGNDATMAGMGCFEGHEGRIEMEPANRWQGGGNCVYDELYF"  
complement(40273..40295)  
/rpt\_family="AT\_rich"  
41836..56291  
/rpt\_family="Oryza sativa gene, repeat sequence Micropon-4  
gi|4586620|dbj|AB010112.1|AB010112"  
complement(41836..41884)  
/rpt\_family="(TA)n"  
41837..56287  
/rpt\_family="Oryza sativa gene, repeat sequence Micropon-4  
gi|4586620|dbj|AB010112.1|AB010112"  
41839..56317  
/rpt\_family="Oryza sativa gene, repeat sequence Micropon-4  
gi|4586620|dbj|AB010112.1|AB010112"  
complement(41839..56317)  
/rpt\_family="Oryza sativa gene, repeat sequence Micropon-4  
gi|4586620|dbj|AB010112.1|AB010112"  
41839..47540  
/rpt\_family="Oryza sativa gene, repeat sequence Micropon-3  
gi|4586621|dbj|AB010113.1|AB010113"  
complement(42556..42578)  
/rpt\_family="AT\_rich"  
complement(43917..43952)  
/rpt\_family="AT\_rich"  
44436..44459  
/rpt\_family="(GA)n"  
complement(44595..44665)  
/rpt\_family="AT\_rich"  
complement(47490..47539)  
/rpt\_family="(TA)n"  
complement(48654..48681)  
/rpt\_family="(GA)n"

```

CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests:
bhan@ncgr.ac.cn
Oryza sativa japonica (nipponbare) genomic DNA, chromosome 4, BAC
clone: OSJNB0020011.
On Jul 12, 2002 this sequence version replaced gi:17998510.
Web site: http://www.ncgr.ac.cn
----- Summary Statistics
Assembly program: phrap
-----
NOTE: This is a PHASE2 sequence. Gaps are shown by nnnnn. Genes
were identified by a combination of several methods: Gene
prediction programs including Egenes (http://www.softberry.com/),
genscan (http://CCK-001.mit.edu/GENSCAN.html), genemarkEMM
(http://genemark.biology.gatech.edu/genemark/), TRNAScan-SE (Sean
Eddy, http://genome.wustl.edu/eddy/TRNAScan-SE/), searches of the
complete sequence against NCBI non redundant protein database (nr)
(http://ncbi.nlm.nih.gov/blast/db) and the EST database at NCGR..
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 88356: contig of 88356 bp in length
88357 88556: gap of 200 bp
88557 113825: contig of 25269 bp in length.

```

QY	843	TACTACCTCCATATTATTAATGATGAAGCGCGTTCACTTTTGTCCACGGTTTGACCATTC	902
Dd	101222	TACTATCTCCATATTTAATGATGATGACGCGTTGACTTTTGTGCCACGTTTGACCATTTC	101163
QY	903	GTCCTATTAAAAAAATTATGCTATCATACTATTATTTATTATTAGTCTGATTCGTCATCA	962
Dd	101162	GTCTTATTCAAAAATTTATGTAATATFCATTTATTTATTTATGCTGATTAGTCATCA	101103
QY	963	AATATTTTTTAAAGCATGCATATAACATTTTCATATTTGCCAAAAAAA-----AAAAG	1015
Dd	101102	AATATCTTTTAGCATATACATATAATTTTTCATATTTGGCACAAAAAATTTGATATTAACG	101043
QY	1016	AATGTCAAATGTTAGTCGAAAGAAGTCATGCTGTCATACATTAAATACGAGGAGTAT	1075
Dd	101042	AATGTCACAAGCTTGGTTGAAAGAATCAACGCGTCATACATTAAATATGAGGGAGTAT	100983
QY	1076	ATGGGATAGCTACAGTAGCAGAGCTATGATATAAAGTAGTAGTATCT	1118
Dd	100982	ATTTTATTGGCTAAAAGGCCCTAAATATATATATTTATATTTT	100940
RESULT	9		
AC	AC092075		
LOCUS		164852 bp	DNA linear
DEFINITION		Oryza sativa chromosome 3 clone OSUnBa0017N12.	*** SEQUENCING IN
ACCESSION		PROGRESS ***,	9 ordered pieces.
VERSION	AC092075.6	GI:20503067	

KEYWORDS	HTG; HGCS_PHASE2.
SOURCE	Oryza sativa.
ORGANISM	Oryza sativa
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE	1 (bases 1 to 164852)
AUTHORS	Buell,C.R., Yuan,Q., Ouyang,S., Liu,J., Tallon,L., Gausberger,K., Kim,M., Overton II,L., Bera,J., Tsitrin,T., Krol,M., Jarrahi,B., Jin,S., Koo,H., Zismann,V., Hsiao,J., Blunt,S., Vanaken,S., Utterbach,T., Feldlyum,T., Yang,Q., Haas,B., Suh,B., Peterson,J., Quackenbush,J., White,O., Salzberg,S. and Fraser,C.
TITLE	Oryza sativa ssp. japonica cv. Nipponbare OSUNBA0017N12 BAC genomic sequence
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 164852)
AUTHORS	Buell,R.
TITLE	Direct Submission
JOURNAL	Submitted (19-JUN-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
REFERENCE	3 (bases 1 to 164852)
AUTHORS	Buell,R.
TITLE	Direct Submission
JOURNAL	Submitted (08-MAY-2002) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
COMMENT	<p>On May 8, 2002 this sequence version replaced gi:18767489.</p> <p>* NOTE: This is a 'working draft' sequence. It currently</p> <p>* consists of 9 contigs. Gaps between the contigs</p> <p>* are represented as runs of N. The order of the pieces</p> <p>* is believed to be correct as given, however the sizes</p> <p>* of the gaps between them are based on estimates that have</p> <p>* provided by the submittor.</p> <p>* This sequence will be replaced</p> <p>* by the finished sequence as soon as it is available and</p> <p>* the accession number will be preserved.</p> <p>* * 1</p> <p>* 10515: contig of 10515 bp in length</p> <p>* 10516</p> <p>* 10615: gap of unknown length</p> <p>* 10616</p> <p>* 23647: contig of 13032 bp in length</p> <p>* 23648</p> <p>* 23747: gap of unknown length</p> <p>* 23748</p> <p>* 36771: contig of 13024 bp in length</p> <p>* 36772</p> <p>* 36871: gap of unknown length</p> <p>* 36872</p> <p>* 44539: contig of 7668 bp in length</p> <p>* 44540</p> <p>* 44639: gap of unknown length</p> <p>* 44640</p> <p>* 72641: contig of 28002 bp in length</p> <p>* 72642</p> <p>* 72741: gap of unknown length</p> <p>* 72742</p> <p>* 111740: contig of 38999 bp in length</p> <p>* 111741</p> <p>* 111840: gap of unknown length</p> <p>* 111841</p> <p>* 142808: contig of 30968 bp in length</p> <p>* 142809</p> <p>* 142808: gap of unknown length</p> <p>* 142909</p> <p>* 149347: contig of 7039 bp in length</p> <p>* 149948</p> <p>* 150047: gap of unknown length</p> <p>* 150048</p> <p>* 164852: contig of 14805 bp in length.</p> <p>Location/Qualifiers</p> <p>1. 164852</p> <p>/organism="Oryza sativa"</p> <p>/cultivar="Nipponbare"</p> <p>/db_xref="taxon:4530"</p> <p>/chromosome="3"</p> <p>/clone="OSUNBA0017N12"</p> <p>/note="japonica cultivar-group"</p>
BASE COUNT	47427 a 35338 c 34787 g 46493 t 807 others
ORIGIN	
Query Match	8.5%; Score 186.8; DB 2; Length 164852;
Best Local Similarity	86.6%; Pred. No. 5e-22;
Matches 220;	Conservative 0; Mismatches 27; Indels 7; Gaps 1;
QY 831	CGTATTACTATTACTACCTCCCATTTTATGATGACGCGGTCACGTTTGTGCGAAC 890
Db 90223	CTTGTTATATATACCATCTCTATATTTTATGATGACGCGGTCACGTTTGTGCGAAC 90282
QY 891	GTTCGACCATTCGCTCTATTATTAATAAAAAAATAATGATCTATTTATTTATTTATTAATGACT 950

Best Local Similarity 85.4%; Pred. No. 5.9e-22;  
Matches 222; Conservative 0; Mismatches 31; Indels 7; Gaps 1;  
QY 826 ATGACCTTATCTATCTACTCCTCATATTTTAACTATGACGCGTTTCTACTTTTGT 885  
Db 132252 ATGTTAGTTTAACTAATCTACTCCTCATATTTTAACTATGACGCGTTTCTACTTTTGT 132311  
QY 886 CCAACCTTTGACCATCTGCTTTATTAATAAATAATATCTATCTATTTATTTATTTAT 945  
Db 132312 CCAACCTTTGACCATCTCTTATTAATAAATAATATCTATCTATTTATTTATTTAT 132371  
QY 946 GACTGATTCGTCATCAATTTTAAAGCATGACATTAATTTTCAATTTTGCACAA 1005  
Db 132372 GACTGATTCGTCATCAATTTTAAAGCATGACATTAATTTTCAATTTTGCACAA 132431  
QY 1006 A-----AAAAACGATGGTCATATGTTAGTTCGAAAGTCATGCTGTCATACATTA 1058  
Db 132432 AATTTTCGAATAAAGCATGACATTAATTTTAAAGCATGACATTAATTTTCAATTTTGCACAA 132491  
QY 1059 AATACGGAGGATTAATG 1078  
Db 132492 AATACGGAGGAGTACTTG 132511

RESULT 11  
CNS07YFX 169663 bp DNA linear HTG 27-JUL-2002  
LOCUS Oryza sativa chromosome 12 clone QJ1618\_C05, \*\*\* SEQUENCING IN  
DEFINITION PROGRESS \*\*\*, in ordered pieces.  
ACCESSION AL713952  
VERSION AL713952.2 GI:22003292  
KEYWORDS HTG; HTGS\_PHASE2; HTGS\_ACTIVEFIN.  
SOURCE Oryza sativa.  
ORGANISM Oryza sativa.  
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.  
1 (bases 1 to 169663)  
AUTHORS Cholsne,N., Orjeda,G., Cattolico,L., Demange,N., Wincker,P.,  
Segures,B., Pelletier,E., Scarpelli,C., Salanoubat,M.,  
Weissenbach,J. and Quetier,F.  
Oryza sativa chromosome 12 sequencing  
2 (bases 1 to 169663)  
Unpublished  
Genoscope.  
Direct Submission  
Submitted (26-JUL-2002) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : segre@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
On Jul 29, 2002 this sequence version replaced gi:19715885.  
IMPORTANT: This sequence is unfinished and does not necessarily  
represent the correct sequence.  
Work on the sequence is in progress and the release of this data is  
based on the understanding that the sequence may change as work  
continue. The sequence may be contaminated with foreign sequence  
from E.coli, yeast, vector, phage, etc.  
The nucleotide sequence of this BAC clone was generated by  
combining Monsanto and Genoscope sequencing data.  
\* NOTE: This is a 'working draft' sequence.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.

FEATURES  
LOCATION/Qualifiers  
SOURCE  
1. .169663  
/organism="Oryza sativa"  
/cultivar="Nipponbare"  
/sub\_species="japonica"  
/db\_xref="taxon:4530"  
/chromosome="12"  
/clone="QJ1618\_C05"  
/clone\_1b="Monsanto"  
BASE COUNT 46256 a 37378 c 36384 g 49645 t  
ORIGIN

Query Match 8.5%; Score 186; DB 2; Length 169663;  
Best Local Similarity 87.2%; Pred. No. 6.9e-22;  
Matches 218; Conservative 0; Mismatches 25; Indels 7; Gaps 1;  
QY 832 CTATTACTTATCTACTCCTCATATTTTAACTATGACGCGTTTCTACTTTTGTCCAAAG 891  
Db 19267 CTACTAGATACTACTCTCCATATTTTAACTATGACGCGTTTCTACTTTTGTCCAAAG 19326  
QY 892 TTGACCACTGCTCTTATTAATAAATAATATCTATCTATTTATTTATTTATTTATTTG 951  
Db 19327 TTGACCACTGCTCTTATTAATAAATAATATCTATCTATTTATTTATTTATTTATTTG 19386  
QY 952 ATTCTGTCATCAATAATTTTAAAGCATGACATTAATTTTCAATTTTGCACAAATAA--- 1009  
Db 19387 ATTCTGTCATCAATAATTTTAAAGCATGACATTAATTTTCAATTTTGCACAAATAA--- 19446  
QY 1010 -----AAAAACGATGGTCATATGTTAGTTCGAAAGTCATGCTGTCATACATTAATAATAC 1064  
Db 19447 TGAATAAATGAATGTCACAAAGTTTGTGAAAGTCACACGCAATCATACATTAATAATAC 19506  
QY 1065 GGAGGGATTA 1074  
Db 19507 GGAGGGAGTA 19516  
RESULT 12  
AX429455/c  
LOCUS AX429455 91552 bp DNA linear PAT 21-JUN-2002  
DEFINITION Sequence 5 from Patent WO0234927.  
ACCESSION AX429455  
VERSION AX429455.1 GI:21540753  
KEYWORDS Oryza sativa.  
SOURCE Oryza sativa.  
ORGANISM Oryza sativa.  
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.  
1  
AUTHORS Farman,M.L., Leong,S.A., Chauhan,R.S. and Durfee,T.J.  
TITLE Plant genes that confer resistance to strains of magnaporthe grisea  
having avr1 co39 cultivar specificity gene  
JOURNAL Patent: WO 0234927-A 5 02-MAY-2002;  
WISCONSIN ALUMNI RES FOUND (US); The Secretary of Agriculture (US)  
FEATURES LOCATION/Qualifiers  
SOURCE  
1. .91552  
/organism="Oryza sativa"  
/db\_xref="taxon:4530"  
BASE COUNT 26676 a 19760 c 17854 g 27262 t  
ORIGIN  
Query Match 8.4%; Score 185.6; DB 6; Length 91552;  
Best Local Similarity 76.2%; Pred. No. 7.3e-22;  
Matches 244; Conservative 0; Mismatches 69; Indels 7; Gaps 1;  
QY 762 TATATAGAGAGATAAATAATTAACAATACTTAACCTGTATATACCCGTTGCAACGGAATAG 821  
Db 33891 TAGATGAGATCACCATAATATTCATAACACCAATTAAGTTGTTTATCATCATGACTCT 33832  
QY 822 TATATGAGACCTTATCTACTACTCCCAATTTTAAATGATGACGCGTTCTACTTT 881  
Db 33831 GGAGATCGCTTGGTACTCAATACTCCCTCTGTATTTTAAATGATGACGTTGTGACTTT 33772  
QY 882 TTGTCACAGTTTGACCATTCGCTTATTAATAAATAATTAATGATCTATCTATTTATTTT 941  
Db 33771 TTGTCACAGTTTGACCATTCGCTTATTAATAAATAATTAATGATCTATCTATTTT 33712  
QY 942 TTATGACTTGATCTCATCAATATTTTAAAGCATGACATTAACATTTTATTTATTTT 1001  
Db 33711 TTATGACTTGATCTCATCAATATTTTAAAGCATGACATTAACATTTTATTTATTTT 33652  
QY 1002 AAAAA-----AAAAACGATGGTCATATGTTAGTTCGAAAGTCATGCTGTCATAC 1054  
ORIGIN

Db 33651 ACARAAATTTGATATAAAGCATGGTCATATGTTGATCGAAAAAGTCACGGCGTGTATAC 33592  
 QY 1055 ATTAAATACGAGGAGATTA 1074  
 Db 33591 GTTAAATACGAGGAGGAGTA 33572

RESULT 13  
 AP004691/c  
 LOCUS  
 DEFINITION Oryza sativa (japonica cultivar-group) chromosome 8 clone P0453D01, linear HTG 21-MAR-2002  
 \*\*\* SEQUENCING IN PROGRESS \*\*\*  
 ACCESSION AP004691.1 GI:18447950  
 VERSION HTG; HTGS\_PHASE2  
 KEYWORDS Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA, clone:P0453D01  
 ORGANISM Oryza sativa (japonica cultivar-group)

REFERENCE  
 AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.  
 TITLE Oryza sativa nipponbare(GA3) genomic DNA, chromosome 8, PAC clone:P0453D01  
 JOURNAL Published only in Database (2002)  
 AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.  
 TITLE Direct Submission  
 JOURNAL Submitted (30-JAN-2002) Takuji Sasaki, National Institute of

Agrobiological Sciences, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
 (E-mail:tsasaki@ias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/, Tel:81-298-38-7441, Fax:81-298-38-7468)  
 NOTE: It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter. This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.  
 \* NOTE: This is a 'working draft' sequence.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.

FEATURES  
 source  
 1. .148441  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /chromosome="8"  
 /clone="P0453D01"  
 BASE COUNT 42350 a 31475 c 33016 g 41500 t 100 others  
 ORIGIN

Query Match 8.4%; Score 185.4; DB 2; Length 148441;  
 Best Local Similarity 83.9%; Pred. No. 8.6e-22;  
 Matches 224; Conservative 0; Mismatches 36; Indels 7; Gaps 1;

QY 833 TATTACTATTATACCTGCATATTTTATGTATGACGCGGTTCACATTTTGTCTCAAGT 892  
 Db 103723 TGTGCTATATACCTGCATATTTTATGTATGACGCGGTTCACATTTTGTCTCAAGT 103664

QY 893 TTGACATCTGCTCTATTAATAAATAATATGTAATCTATTATTTATTTATGACTTGA 952  
 Db 103663 TTGACATCTGCTCTATTAATAAATAATATGTAATCTATTATTTATTTATGACTTGA 103604

QY 953 TTGCTCATCAATAATTTTAAAGCATGACATAAATTTTCATATTTGCAAAA----- 1006  
 Db 103603 TTCTTCATCAATAATTTTAAAGCATGACATAAATTTTCATATTTGCAAAAATTTT 103544

QY 1007 -AAAAACGAATGGTCAATGTTAGTGAAGTCAATGGTGCATACATTAATAATACG 1065  
 Db 103543 GAATAAACGAATGGTCAATGTTAGTGAAGTCAATGGTGCATACATTAATAATACG 103484

QY 1066 GAGGATTATATCGGATACGTACAGTA 1092  
 Db 103483 GAGGAGTATGTTTTCATGATTA 103457

RESULT 14  
 AC120885  
 LOCUS  
 DEFINITION Oryza sativa (japonica cultivar-group) chromosome 11 clone Ba0042J05, linear HTG 13-MAY-2002  
 \*\*\* SEQUENCING IN PROGRESS \*\*\*  
 ACCESSION AC120885.1 GI:20531970  
 VERSION HTG; HTGS\_PHASE2  
 KEYWORDS Oryza sativa (japonica cultivar-group)  
 SOURCE Oryza sativa (japonica cultivar-group)  
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE  
 AUTHORS Khurana,J.P., #Linton,E.W., #Messing,J., Mohanty,A., Bharti,A.K., Raghuvanshi,S., Khurana,P. and Tyagi,A.K.  
 TITLE Oryza sativa (japonica cultivar-group) chromosome 11 BAC clone Ba0042J05, sequencing in progress  
 JOURNAL Unpublished  
 REMARK # The Plant Genome Initiative at Rutgers - Waksman Institute, Rutgers University 190 Frelinghuysen Road, Piscataway, NJ 08873, USA

REFERENCE  
 AUTHORS Khurana,J.P., #Linton,E.W., #Messing,J., Mohanty,A., Bharti,A.K., Raghuvanshi,S., Khurana,P. and Tyagi,A.K.  
 TITLE Direct Submission  
 JOURNAL Submitted (13-MAY-2002) Indian Initiative for Rice Genome Sequencing, Department of Plant Molecular Biology, University of Delhi South Campus, New Delhi, Delhi 110021, India  
 REMARK # The Plant Genome Initiative at Rutgers - Waksman Institute, Rutgers University 190 Frelinghuysen Road, Piscataway, NJ 08873, USA

COMMENT  
 \* NOTE: This is a 'working draft' sequence. It currently consists of 7 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.

1  
 1207 1206: contig of 1206 bp in length  
 1307 1306: gap of unknown length  
 21093 21093: contig of 19787 bp in length  
 21094 21193: gap of unknown length  
 21194 64937: contig of 43744 bp in length  
 64938 65037: gap of unknown length  
 65038 81321: contig of 16284 bp in length  
 81322 81421: gap of unknown length  
 113858 113857: contig of 32436 bp in length  
 113958 120340: contig of 6283 bp in length  
 120341 120340: gap of unknown length  
 120341 154494: contig of 34154 bp in length.  
 Location/Qualifiers  
 1. .154494

/organism="Oryza sativa (japonica cultivar-group)"  
 /cultivar="Nipponbare"  
 /sub\_species="japonica cultivar-group"  
 /db\_xref="taxon:39947"  
 /chromosome="11"  
 /clone="Ba0042J05"  
 BASE COUNT 44561 a 32837 c 31980 g 44505 t 611 others  
 ORIGIN

Query Match 8.4%; Score 185.2; DB 2; Length 154494;  
 Best Local Similarity 82.0%; Pred. No. 9.4e-22;



```
Matches 228; Conservative 0; Mismatches 43; Indels 7; Gaps 1;
QY 807 GTTCACGGAATAGTATATGGACCTATTACTATTACTACCTCCATATTTTAATGTAT 866
|||||
Db 144949 GATGATAGGCAAGTGAAGCAGCTGCTTTATTTTATTACTACCTCCGATTATTAATGCAT 145008
|||||
QY 867 GACGCGCTTCACTTTTGTCCAAAGTTTGACCACTTGTCTTATTAATAAAAAAATATGTAT 926
|||||
Db 145009 GACCTGCTTACCTTTTCTCTAAGTTTGACCATTCGCTCTATTCAAAAATTTATGTAA 145068
|||||
QY 927 CTATTATTATTTTATATGACTTGATCTGTCATCAATATTTTAAAGCATGACATAAA 986
|||||
Db 145069 TTATCATTTTATTTATGACTTGATTTGTTTATCAATATTTCTTTAAGCATGACATAAA 145128
|||||
QY 987 CATTTTCATATTTGCAAAA-----AAAAAAGCAATGGTCAATCTTAGTCGAAAG 1039
|||||
Db 145129 TATTTTCATATTTGCAAAAATTTTGAATAAACGATGGTCAACAACTTGGTCGAAAG 145188
|||||
QY 1040 TCAATGGTGTCAATATAAATACGGAGGATATAT 1077
|||||
Db 145189 TCAACGGCATCATACATATAAATACGGAGGAGTATAT 145226
|||||

RESULT 15
AP005305/c
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) chromosome 8 clone P0024C06,
*** SEQUENCING IN PROGRESS ***, 3 ordered pieces.
ACCESSION
AP005305
VERSION
AP005305.1 GI:21280335
KEYWORDS
HTG; HTGS_PHASE2.
SOURCE
Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
clone:P0024C06.
ORGANISM
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1
Sasaki,T., Matsumoto,T. and Katayose,Y.
Oryza sativa nipponbare(GN3) genomic DNA, chromosome 8, PAC
clone:P0024C06
Published Only in Database (2002)
REFERENCE
2 (bases 1 to 179486)
Sasaki,T., Matsumoto,T. and Katayose,Y.
Direct Submission
Submitted (29-MAY-2002) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rpg.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOPE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 22904: contig of 22904 bp in length
* 22905 23080: gap of 176 bp
* 23081 23081: contig of 1 bp in length
* 23082 23186: gap of 105 bp
* 23187 179486: contig of 156300 bp in length.
FEATURES
Location/Qualifiers
1..179486
/organism="Oryza sativa (japonica cultivar-group)"
```

```
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="8"
/clones="P0024C06"
BASE COUNT 50872 a 39291 c 39711 g 49023 t 589 others
ORIGIN
Query Match 8.4%; Score 184.6; DB 2; Length 179486;
Best Local Similarity 85.9%; Pred.No.1.2e-24;
Matches 219; Conservative 0; Mismatches 29; Indels 7; Gaps 1;
QY 834 ATTACTTATTACTACTCCATATTTTAAATGATGACGCGGTTCACTTTTGTCCAAAGTT 893
|||||
Db 15016 ATTATTTTGTACTCCCTCCATATTTTAAATGATGACGCGGTTGACTTTTATCCAAATAT 14957
|||||
QY 894 TGACCAATCGTCTTATTAATAAAAAAATTAATGATATCTATTTATTTATTTATGACTTGTAT 953
|||||
Db 14956 TGACCAATCGTCTTATTAATAAAAAAATTTATGATATCTATTTATTTATTTATGACTTGTAT 14897
|||||
QY 954 TCGTCATCAATATTTTAAAGCATGACATAAAACATTTTCATATTTTGCAAAAAA--- 1009
|||||
Db 14896 TTATCATCAAAATATTTCTTTAAGCATGACATAAAATATCTCATATTTGCACAAAAAATTTG 14837
|||||
QY 1010 ---AAAAGCAATGGTCAATATTTAGTTCGAAAAAGTCAATGGTGCATACATTTAAATAACGG 1066
|||||
Db 14836 AATAAAATGAATGATCAAAACGTTGGTCAAAAAAGTCAACGACGTCATACATTAATAAATACGG 14777
|||||
QY 1067 AGGCATATATATGGGA 1081
|||||
Db 14776 AGGAGATATCTAGAA 14762
|||||
```

Search completed: March 26, 2003, 14:00:53  
Job time : 7637.65 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2003, 18:27:39 ; Search time 335.492 Seconds  
(without alignments)  
14774.266 Million cell updates/sec

Title: US-09-702-134-7212\_COPY\_10200\_12400  
Perfect score: 2201  
Sequence: 1 cgtcgttggtttgtgcatc.....tagcttggtttgttgta 2201

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002.\*  
1: /SID32/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SID32/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SID32/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SID32/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SID32/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SID32/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SID32/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SID32/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SID32/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SID32/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SID32/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SID32/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SID32/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SID32/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SID32/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SID32/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SID32/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SID32/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SID32/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SID32/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SID32/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SID32/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SID32/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SID32/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	185.6	8.4	91552	24	AAD38803
2	125.2	5.7	807	21	AAC52171
c 3	124.4	5.7	356	24	ABQ85736
4	124.4	5.7	1126	21	AAC47067
5	111.2	5.1	4569	24	ABK15663
6	110.2	5.0	689	24	ABQ85731
7	106.6	4.8	5241	21	AAQ99467
8	101	4.6	7175	19	AAV35027
c 9	100	4.5	33675	24	AAD31202

c 10	100	4.5	69300	24	AAD38804
c 11	98.6	4.5	4569	24	ABK15663
12	98.2	4.5	91552	24	AAD38803
13	95.4	4.3	694	21	AAC48763
14	95.4	4.3	695	21	AAQ44265
15	91.4	4.2	1591	18	AAQ93780
c 16	90.8	4.1	6901	15	AAQ73703
17	84.6	3.8	1032	21	AAC44276
c 18	80.8	3.7	2317	20	AAQ33532
c 19	80.6	3.7	5241	21	AAQ99467
20	74.6	3.4	1978	20	AAQ32597
21	72.6	3.3	33675	24	AAD31202
c 22	70.2	3.2	7175	19	AAV35027
c 23	68.4	3.1	1978	20	AAQ32597
c 24	68	3.1	1428	21	AAQ99466
c 25	66	3.0	4371	21	AAQ99469
26	66	3.0	8305	24	ABL33569
27	64.6	2.9	13326	24	ABL33712
28	64.4	2.9	6901	15	AAQ73703
c 29	62	2.8	17953	24	AAD38802
c 30	61.8	2.8	1591	18	AAQ93780
31	61.4	2.8	5989	24	ABL54319
32	61.4	2.8	9810	24	ABL32427
33	61.2	2.8	50000	24	ABL55644
34	60.8	2.8	12120	24	AAQ96695
c 35	60.6	2.8	2699	22	AAQ27620
c 36	60.6	2.8	2784	22	AAQ27619
c 37	60.6	2.8	4077	22	AAQ27621
c 38	60.4	2.7	32392	24	ABL56203
39	60.2	2.7	21313	22	AAQ82710
c 40	60.2	2.7	222930	24	ABK84349
c 41	60	2.7	17183	24	ABL32486
c 42	60	2.7	18598	24	ABL32387
c 43	59.8	2.7	4590	7	AAQ60472
c 44	59.8	2.7	109906	24	ABK94411
c 45	59.6	2.7	550	19	AAV63932

# ALIGNMENTS

RESULT 1  
AAD38803/c  
ID AAD38803 standard; DNA; 91552 BP.  
XX AAD38803;  
AC AAD38803;  
XX  
DT 23-SEP-2002 (first entry)  
XX  
DE BAC clone K6P36 from rice variety CO39.  
XX  
KW Rice; chromosome 11; Indica rice cultivar; CO39; avirulence gene;  
KW AVRI-CO39; rice blast pathogen; transgenic plant; plant breeding;  
KW resistance; agricultural; horticultural; plant protectant; ds.  
XX  
OS Oryza sativa.  
XX  
PN WO200234927-A2.  
XX  
PD 02-MAY-2002.  
XX  
PF 19-OCT-2001; 2001WO-US46331.  
XX  
PR 20-OCT-2000; 2000US-242313P.  
PR 09-JUL-2001; 2001US-303897P.  
XX  
PA (WISC ) WISCONSIN ALUMNI RES FOUND.  
PA (USDA ) US DEPT OF AGRICULTURE.  
XX  
PI Leong SA, Farman ML, Chauhan RS, Durfee TJ;  
XX WPI; 2002-471442/50.  
XX

PT New PIC039(t) polynucleotides from chromosome 11 of Indica rice  
 PT cultivar CO39 useful for conferring or improving resistance of plants  
 PT to strains of Magnaporthe grisea or other plant pathogens with  
 PS avirulence gene AVR1-CO39 -  
 XX Example 2; Page 68-110; 175pp; English.  
 XX  
 CC The invention relates to a polynucleotide isolated from chromosome 11 of  
 CC Indica rice cultivar CO39, flanked by marker R2316 and RG1094 comprising  
 CC one or more genes that confer resistance to strains of Magnaporthe grisea  
 CC having avirulence gene AVR1-CO39. The Pi-CO39(t) polynucleotides are  
 CC useful for conferring or improving resistance of plants to strains of the  
 CC rice blast pathogen, Magnaporthe grisea and other pathogens. The  
 CC Pi-CO39(t) nucleic acids may be used as probes to detect the presence of  
 CC and/or expression of Pi-CO39(t) genes, and to produce large quantities of  
 CC pure Pi-CO39(t) proteins. Purified gene products of Pi-CO39(t) are useful  
 CC in producing polyclonal or monoclonal antibodies useful as sensitive  
 CC detection reagents for the presence and accumulation of Pi-CO39(t)  
 CC polypeptides. The transgenic plants are useful for plant breeding or  
 CC directly in agricultural or horticultural applications. The present  
 CC sequence is BAC clone K6P36 from rice variety CO39.  
 XX  
 SQ Sequence 91552 BP; 26676 A; 19760 C; 17854 G; 27262 T; 0 other;  
 Query Match 8.4%; Score 185.6; DB 24; Length 91552;  
 Best Local Similarity 76.2%; Pred. No. 7.5e-30;  
 Matches 244; Conservative 0; Mismatches 69; Indels 7; Gaps 1;  
 QY 762 TATCTAGAGAGATTAAATAATTATACAAACTGATGATACCGTGGCAACGGAATAG 821  
 DB 33891 TAGATGAGATCAACCAATATTTTCATCAACCAATTAAGTGTCTATACATGTCTGT 33832  
 QY 822 TAATATGGACATTAATCTATTAATCTCCATATTTAATGATGACGCGGTCTACTTT 881  
 DB 33831 GGAGATCGCTGGTACTACATATCTCCTCTGATTTTAAGTATGACGCTGTGTGACTTT 33772  
 QY 882 TTGTCACGCTTGACCATCTGCTTTATTAATAAATAATGATATCATATTTATTTATTA 941  
 DB 33771 TTGTTCAACGCTTGACCATCTGCTTTATTAATAAATAATGATATCATATTTATTTA 33712  
 QY 942 TTATGACTTGATCTGTCATCAATATTTTAAAGCATGACATCAATCTTCAATTTTTC 1001  
 DB 33711 TTATGACTTGATCTGTCATCAATATTTTAAAGCATGACATCAATCTTCAATTTTTC 33652  
 QY 1002 AAAAA-----AAAAAACGAATGGTCAAAATGTTAGTCGAAAAGTCAATGGTGCATAC 1054  
 DB 33651 ACRAAAATTTGATAAATAAACGAATGGTCAAAATGTTAGTCGAAAAGTCAACGGGTCATAC 33592  
 QY 1055 ATTAATAATACGAGGGATTA 1074  
 DB 33591 GTTAAATAACGAGGGAGTA 33572  
 RESULT 2  
 AAC52171  
 ID AAC52171 standard; DNA; 807 BP.  
 XX  
 AC AAC52171;  
 XX  
 DT 18-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 70520.  
 XX  
 KW Hybridisation assay; genetic mapping; gene expression control;  
 KW protein identification; signal transduction pathway;  
 KW metabolic pathway; promoter; termination sequence; ss.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN EP1033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX

PF 25-FEB-2000; 2000EP-0301439.  
 XX  
 PR 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128234.  
 PR 08-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129845.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 23-APR-1999; 99US-0130891.  
 PR 28-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.  
 PR 30-APR-1999; 99US-0132407.  
 PR 04-MAY-1999; 99US-0132484.  
 PR 05-MAY-1999; 99US-0132485.  
 PR 06-MAY-1999; 99US-0132486.  
 PR 06-MAY-1999; 99US-0132487.  
 PR 07-MAY-1999; 99US-0132863.  
 PR 11-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 14-MAY-1999; 99US-0134370.  
 PR 18-MAY-1999; 99US-0134768.  
 PR 19-MAY-1999; 99US-0134941.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135353.  
 PR 24-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136392.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139492.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139763.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 22-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 23-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140695.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 01-JUL-1999; 99US-0142154.  
 PR 02-JUL-1999; 99US-0142055.  
 PR 06-JUL-1999; 99US-0142390.  
 PR 08-JUL-1999; 99US-0142803.

PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0158393.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
Query Match 5.7%; Score 125.2; DB 21; Length 807;		
Best Local Similarity 65.9%; Pred. No. 1.5e-17;		
Matches 178; Conservative 2; Mismatches 90; Indels 0; Gaps		
QY	1214	GAGRAGCTGAGGAGCACAATCGAGGAGGCGGGAGACGAGTGGCGAAGCAGCCAG 1273
Db	379	GAGCCAAATTAGCAGCATGTTGAGAGAGAGTAGGCAAACTCTTCAAGAAACACAGTCAT 438
QY	1274	CTTGTGAAGGAGGTGGAGCTGGCCCTCCGCGCGCGGTGGCGAGCTCAGCCGAGGTCC 1333
Db	439	CKTGTGAGAGAGTTGATTAAGACTCTCTGTCGTGGTGGAGACTTTGGTAAAGGCCCT 498
QY	1334	AGATTTGCGAGTGGCAGATCACTCTCTTCACCAAGCGCACCGGTCATCCGGCGCGAG 1393
Db	499	AGGATTCGAAGATGTGAGGTGACATGTTTACAAGAAGCATGGTGTGTCCGTGCTGAG 558
QY	1394	GAGGAGCCGAGTCCACTACGCGACATCGACTCCGCTCCPCCATCATCAAGAGGAAG 1453
Db	559	GAGATGCTGAGCAGATATACCTGTATCGACTTGGTATCAACAGTATATACAGAGGAAG 618
QY	1454	CTCAGAGATCATAGGAGAGGAGCCGAC 1483
Db	619	CTGAGGAGATCAAGGAGAGGACTCAGAC 648

RESULT 3	
ABQ85736/c	
ID	ABQ85736 standard; DNA; 356 bp.
XX	
AC	ABQ85736;
XX	
DT	05-SEP-2002 (first entry)
XX	
DE	<i>Arabidopsis thaliana</i> expressed polynucleotide SEQ ID NO 606.
XX	
XX	Plant; <i>Arabidopsis</i> ; transgenic; fungicide; insecticide; ds.

```
XX Arabidopsis thaliana.
XX OS
XX US2002062014-A1.
XX PD
XX 23-MAY-2002.
XX PF
XX 26-JAN-2001; 2001US-0770791.
XX PR
XX 27-JAN-2000; 2000US-178480P.
XX PA
XX (GORL/) GORLACH J.
XX PA
XX (ANYI/) AN Y.
XX PA
XX (HAMI/) HAMILTON C M.
XX PA
XX (PRIC/) PRICE J L.
XX PA
XX (RAIN/) RAINES T M.
XX PA
XX (YOYI/) YU Y.
XX PA
XX (RAME/) RAMEAKA J G.
XX PA
XX (PAGE/) PAGE A.
XX PA
XX (MATH/) MATHIEU A V.
XX PA
XX (LEDF/) LEDFORD B L.
XX PA
XX (WOES/) WOESSNER J P.
XX PA
XX (HAAS/) HAAS W D.
XX PA
XX (GARC/) GARCIA C A.
XX PA
XX (KRIC/) KRICKER M.
XX PA
XX (SLAT/) SLATER T.
XX PA
XX (DAVI/) DAVIS K R.
XX PA
XX (ALLE/) ALLEN K.
XX PA
XX (HOFF/) HOFFMAN N.
XX PA
XX (HURB/) HURBAN P.
XX PI
XX Goriach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;
XX PI Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
XX PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;
XX PI Hurban P;
XX XX
XX WPI; 2002-479265/51.
XX PT
XX New nucleic acid sequences of Arabidopsis thaliana and their encoded
XX PT products are useful to produce transgenic plants, to screen for
XX PT biologically active agents such as fungicides and insecticides and in
XX PT genetic studies -
XX PS
XX Claim 1; SEQ ID NO 606; 18pp + Sequence Listing; English.
XX CC
XX The invention relates to a novel nucleic acid of Arabidopsis thaliana
XX CC comprising a sequence capable of hybridising under stringency to one of
XX CC the 999 sequences referred to but not defined in the specification
XX CC (AB085131-AB086129). The nucleic acid sequences are useful to identify
XX CC homologous or related genes, to produce compositions that modulate
XX CC expression or function of the encoded protein, to map functional regions
XX CC of the protein, to study associated physiological pathways, to
XX CC genetically manipulate cells and plants. The encoded products are useful
XX CC to screen for biologically active agents such as fungicides or
XX CC insecticides and to elucidate biochemical pathways.
XX XX
XX SQ Sequence 356 BP; 83 A; 103 C; 62 G; 108 T; 0 other;
Query Match 5.7%; Score 124.4; DB 24; Length 356;
Best Local Similarity 66.3%; Pred. No. 1.6e-17;
Matches 179; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
QY 1214 GAGAGGTGTAAGGAGGACATCGAGGAGAGGCGGGGAGAGCAGTGGCGGACGAGCGGAG 1273
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
284 GAGCCAAATTAAGCAGCATGTTGAGAGGAGTAGGCAAAATCTGTTACAGAAACACAGTCAT 225
QY 1274 CTTGTGAGGAGGTGGACGTGGCGCTCTCCCGCGGGTGGAGCTCAGCGGAGTCGC 1333
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
224 CTTGTGAGAGAGTGTATGATGTAAGACTCTCTGTTGGTGGAGAGTGTGGTAAGGCCCT 165
QY 1334 AAGATTTTCAGGTGCGAGATCACTCTCTTACCAAGCGCCGCGGTATCCGCGCGGAG 1393
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
164 AGGATTCGAAGATGTGAGGTGACATGTTTACAAAGAGGACATGGTGTGCTGCTGAG 105
```

```
QY 1394 GAGGAGCGCGAGTCCACCTACGCCAGCATCGACCTCGCTCCCTCCATCATCAAGAGGAG 1453
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
104 GAGAGTGTGAGAGCAGTATACGCTTGTATCGACTTGGTATCAACAGGATATACAGAGGAG 45
QY 1454 CTCAGAGATCAAGGAGAGGAGGAGGACCCGAC 1483
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
44 CTGAGGAAGATCAAGGAGGAGGAGGACTCAGAC 15
RESULT 4
BAC47067
ID ARC47067 standard; DNA; 1126 BP.
XX AC
XX AAC47067;
XX DT
XX 18-OCT-2000 (first entry)
XX DE
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 52441.
XX KW
XX Hybridisation assay; genetic mapping; gene expression control;
XX KW protein identification; signal transduction pathway;
XX KW metabolic pathway; promoter; termination sequence; ss.
XX OS
XX Arabidopsis thaliana.
XX FN
XX EP1033405-A2.
XX PD
XX 06-SEP-2000.
XX PF
XX 25-FEB-2000; 2000EP-0301439.
XX PR
XX 25-FEB-1999; 99US-0121825.
XX PR
XX 03-MAR-1999; 99US-0123180.
XX PR
XX 09-MAR-1999; 99US-0123548.
XX PR
XX 23-MAR-1999; 99US-0125788.
XX PR
XX 25-MAR-1999; 99US-0126264.
XX PR
XX 29-MAR-1999; 99US-0126785.
XX PR
XX 01-APR-1999; 99US-0127462.
XX PR
XX 06-APR-1999; 99US-0128334.
XX PR
XX 08-APR-1999; 99US-0128714.
XX PR
XX 16-APR-1999; 99US-0129845.
XX PR
XX 19-APR-1999; 99US-0130077.
XX PR
XX 21-APR-1999; 99US-0130449.
XX PR
XX 23-APR-1999; 99US-0130510.
XX PR
XX 23-APR-1999; 99US-0130891.
XX PR
XX 28-APR-1999; 99US-0131449.
XX PR
XX 30-APR-1999; 99US-0132048.
XX PR
XX 30-APR-1999; 99US-0132407.
XX PR
XX 04-MAY-1999; 99US-0132484.
XX PR
XX 05-MAY-1999; 99US-0132485.
XX PR
XX 06-MAY-1999; 99US-0132486.
XX PR
XX 07-MAY-1999; 99US-0132487.
XX PR
XX 11-MAY-1999; 99US-0132863.
XX PR
XX 14-MAY-1999; 99US-0134256.
XX PR
XX 14-MAY-1999; 99US-0134219.
XX PR
XX 14-MAY-1999; 99US-0134221.
XX PR
XX 18-MAY-1999; 99US-0134370.
XX PR
XX 19-MAY-1999; 99US-0134941.
XX PR
XX 20-MAY-1999; 99US-0135124.
XX PR
XX 21-MAY-1999; 99US-0135353.
XX PR
XX 24-MAY-1999; 99US-0135629.
XX PR
XX 25-MAY-1999; 99US-0136021.
XX PR
XX 27-MAY-1999; 99US-0136392.
XX PR
XX 28-MAY-1999; 99US-0136782.
XX PR
XX 01-JUN-1999; 99US-0137222.
XX PR
XX 03-JUN-1999; 99US-0137528.
XX PR
XX 04-JUN-1999; 99US-0137502.
XX PR
XX 07-JUN-1999; 99US-0137724.
XX PR
XX 08-JUN-1999; 99US-0138094.
XX PR
XX 10-JUN-1999; 99US-0138540.
```



Db 381 GAGCAATTAAGCAGCATGTTGAAGAAAGTAGGCAAAATCTGTTCAGAAACACAGTCAI 440  
 QY 1274 CTGTGGAAGAGGTGGACGTGCGCTCTCCCGCGGTGGCGAGCTCAGCGAGTCCC 1333  
 Db 441 CTGTGGAAGAGGTGGACGTGCGCTCTCCCGCGGTGGCGAGCTCAGCGAGTCCC 500  
 QY 1334 AAGATTGCGAGTGGAGATCACTCTCTACCAAGCGCCAGCGCTCATCGCCCGCGAG 1393  
 Db 501 AGGATTCGAAGATGGAGTGGACATTTTACAAAGAGCATGTTGTGCGTCTCTGAG 560  
 QY 1394 GAGGAGCGGAGTCCACCTACCGCCAGCATCGACCTCGCTCTCTCTCTCTCTCTCT 1453  
 Db 561 GAGATGCTGAGCAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 620  
 QY 1454 CTGAGGAGATCAAGGAGAGGAGCGAC 1483  
 Db 621 CTGAGGAGATCAAGGAGAGGAGCTGAGAC 550

RESULT 5  
 ABK15663  
 ID ABK15663 standard; DNA; 4569 BP.  
 XX  
 AC ABK15663;  
 XX  
 DT 21-MAY-2002 (first entry)  
 XX  
 DE Rice lipoxigenase gene 4.5kb fragment.  
 XX  
 KW Rice; ds; lipoxigenase; RCI-1; transgenic; plant; plant antifungal;  
 KW rice chemically induced cDNA; promoter; trans peptide; plastid;  
 KW fungal mycotoxin inhibitor; plant breeding; pBSK+LOX4a.  
 XX

OS Oryza sativa.  
 XX  
 XX W0200206490-AI.  
 XX  
 PD 24-JAN-2002.  
 XX  
 XX 12-JUL-2001; 2001WO-EP08085.  
 XX  
 PR 13-JUL-2000; 2000GB-0017275.  
 PR 15-SEP-2000; 2000GB-002739.  
 XX

XX (SYGN ) SYNGENIA PARTICIPATIONS AG.  
 PA (UYZU-) UNIV ZUERICH.  
 XX  
 XX Dudley R, Schaffrath, Lawton KA;  
 XX  
 XX WPI; 2002-188550/24.  
 DR  
 XX Novel isolated nucleic acid encoding a promoter which is capable of  
 PT driving chemically inducible but not wound- or pathogen-inducible  
 PT expression of an associated nucleotide sequence  
 XX  
 PS Claim 3; Page 68-69; 88pp; English.

XX The invention relates to an isolated nucleic acid molecule (a promoter of  
 CC rice chemically induced cDNA (RCI-1), which encodes a lipoxigenase)  
 CC capable of driving chemically-inducible but not wound- or pathogen-  
 CC inducible expression of an associated nucleotide sequence. Also  
 CC included are the RCI-1 cDNA, its encoded protein, a 4.5kb genomic clone  
 CC for the lipoxigenase gene, promoter fragments, the lipoxigenase transit  
 CC peptide which directs expressed proteins to the plastid, a vector  
 CC comprising the promoter or fragments and a transgenic plant comprising  
 CC the vector. The promoter or fragments are useful for expressing a  
 CC nucleotide sequence of interest. The transit peptide is useful for  
 CC targeting an associated protein of interest to plastids. A nucleic acid  
 CC which expresses polypeptide having lipoxigenase activity is useful for  
 CC inhibiting fungal mycotoxins when transformed into a plant. The  
 CC lipoxigenase is useful for inhibiting fungal mycotoxins. The promoter is  
 CC useful for regulating transcription of a chemically inducible but not  
 CC wound or pathogen inducible gene, which involves applying a chemical

CC regulator to a plant or seed containing a chemically regulatable  
 CC nucleotide sequence. Transgenic plants as described above are useful for  
 CC breeding improved plant lines that for example increase the effectiveness  
 CC of conventional methods such as herbicide or pesticide treatment or allow  
 CC to dispense with the methods due to their modified genetic properties.  
 CC New crops with improved stress tolerance can be obtained that, due to  
 CC their optimised genetic equipment yield harvested product of better  
 CC quality than products that were not able to tolerate comparable adverse  
 CC developmental conditions. The present sequence is the 4.5kb  
 CC fragment of the RCI-1 gene from plasmid pBSK+LOX4a.

XX  
 SQ Sequence 4569 BP; 1234 A; 1051 C; 980 G; 1304 T; 0 other;  
 Query Match 5.1%; Score 111.2; DB 24; Length 4569;  
 Best Local Similarity 68.4%; Pred. No. 3.7e-14;  
 Matches 186; Conservative 0; Mismatches 78; Indels 8; Gaps 2;  
 QY 843 TACTACCTCCATATTTAATGATGACGCGCTGACCTTTTGTTCACACCTTTGACCATTC 902  
 Db 2845 TACTACCTCTGTTTTTAAATAGATGACGCGCTGACCTTTTCTCACATGTTGACCATTC 2904  
 QY 903 GTCTTATTAATAAATAATGATGATATATTTTATTTATGACTTGTATGCTGTCATCA 962  
 Db 2905 GTCTTATTTTAAATAATATATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 2962  
 QY 963 AATATTTTAAAGCATGACATAAACAATTTTCATATTTTCAAAAAA-----AAAAAACA 1016  
 Db 2963 ATAGTACTTTAAGTGTGATTTATATCTTATATATTTTATATTTTGAATAAGACA 3022  
 QY 1017 ATGGTCAATTTAGTGTGCAAAAGTCAATGCTCATACATTAATAACGGAGGATTATA 1076  
 Db 3023 ATGGTCAATTTAGTGTGCAAAAGTCAATGCTCATACATTAATAACGGAGGATTATA 3082  
 QY 1077 TGGGATAGCTACAGTACAGATCATGATAAA 1108  
 Db 3083 TGGTATCATCGGAATTAATAATAGACGCTATA 3114

RESULT 6  
 ABQ65731  
 ID ABQ65731 standard; DNA; 689 BP.  
 XX  
 AC ABQ65731;  
 XX  
 DT 21-AUG-2002 (first entry)  
 XX  
 DE Arabidopsis thaliana polynucleotide SEQ ID NO 308.  
 XX  
 KW Arabidopsis thaliana; thale cress; plant; transgenic; GMO; disease;  
 KW stress; metabolic pathway; biosynthetic pathway; nutrition; fungicide;  
 KW insecticide; antibiotic; ds.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 EN US2002059663-AI.  
 XX  
 PD 16-MAY-2002.  
 XX  
 PF 26-JAN-2001; 2001US-0770149.  
 XX  
 PR 27-JAN-2000; 2000US-178506P.  
 XX  
 XX (GORL/) GORLACH J.  
 PA (ANYI/) AN Y.  
 PA (HAMI/) HAMILTON C M.  
 PA (PRIC/) PRICE J L.  
 PA (RAIN/) RAINES T M.  
 PA (YUYI/) YU Y.  
 PA (RAME/) RAMEAKA J G.  
 PA (PAGE/) PAGE A.  
 PA (MATH/) MATHW A V.  
 PA (LEDE/) LEDFORD B L.  
 PA (WOES/) WOESSNER J P.









Db 65409 TTATGTATATATATATATAGTTTTACATPATTTCAAAAAGTTTTTAAATAAAACGAACGG 65350

QY 1021 TCAAATGTTAGTCGAAAAGTCAATCGTGTCAATATAAATACGAGAGGATTATA 1076

Db 65349 TTAACTATATTTTAAAAAGTCAACGAGCTCAATATTTTTGGACACGAGGAGATATA 65294

RESULT 11	
ABK15663/c	
ID	ABK15663 standard; DNA; 4569 BP.
XX	
XX	ABK15663;
XX	
DT	21-MAY-2002 (first entry)
XX	
DE	Rice lipoxigenase gene 4.5kb fragment.
XX	
XX	Rice; ds; lipoxigenase; RCI-1; transgenic; plant; plant antifungal;
KW	rice chemically induced cDNA; promoter; transit peptide; plastid;
KW	fungal mycotoxin inhibitor; plant breeding; pBSK+LOX4a.
XX	
OS	Oryza sativa.
XX	
PN	WO200206490-A1.
XX	
XX	24-JAN-2002.
XX	
XX	12-JUL-2001; 2001WO-EP08085.
PF	
XX	
PR	13-JUL-2000; 2000GB-0017275.
XX	
PR	15-SEP-2000; 2000GB-0022739.
XX	
XX	(SYGN ) SYNGENTA PARTICIPATIONS AG.
PA	(UYZU-) UNIV ZUERICH.
PA	
PA	Dudler R, Schafrath , Lawton KA;
PI	
XX	WPI; 2002-188550/24.
DR	
XX	
XX	Novel isolated nucleic acid encoding a promoter which is capable of
PT	driving chemically inducible but not wound- or pathogen-inducible
PT	expression of an associated nucleotide sequence
PT	
XX	
XX	Claim 3; Page 68-69; 88pp; English.
PS	
XX	
XX	The invention relates to an isolated nucleic acid molecule (a promoter of
CC	rice chemically induced cDNA (RCI-1), which encodes a lipoxigenase)
CC	capable of driving chemically-inducible but not wound- or pathogen-
CC	inducible expression of an associated nucleotide sequence. Also
CC	included are the RCI-1 cDNA, its encoded protein, a 4.5kb genomic clone
CC	for the lipoxigenase gene, promoter fragments, the lipoxigenase transit
CC	peptide which directs expressed proteins to the plastid, a vector
CC	comprising the promoter or fragments and a transgenic plant comprising
CC	the vector. The promoter or fragments are useful for expressing a
CC	nucleotide sequence of interest. The transit peptide is useful for
CC	targeting an associated protein of interest to plastids. A nucleic acid
CC	which expresses polypeptide having lipoxigenase activity is useful for
CC	inhibiting fungal mycotoxins when transformed into a plant. The
CC	lipoxigenase is useful for inhibiting fungal mycotoxins. The promoter is
CC	useful for regulating transcription of a chemically inducible but not
CC	wound or pathogen inducible gene, which involves applying a chemical
CC	regulator to a plant or seed containing a chemically regulatable
CC	nucleotide sequence. Transgenic plants as described above are useful for
CC	breeding improved plant lines that for example increase the effectiveness
CC	of conventional methods such as herbicide or pesticide treatment or allow
CC	to dispense with the methods due to their modified genetic properties.
CC	New crops with improved stress tolerance can be obtained that, due to
CC	their optimised genetic equipment yield harvested product of better
CC	quality than products that were not able to tolerate comparable adverse
CC	developmental conditions. The present sequence is the 4.5kb
CC	fragment of the RCI-1 gene from plasmid pBSK+LOX4a.
XX	
SQ	Sequence 4569 BP; 1234 A; 1051 C; 980 G; 1304 T; 0 other;

	Query Match	4.5%	Score 98.6;	DB 24;	Length 4569;
	Best Local Similarity	61.8%;	Pred. No. 2.le-11;		
	Matches 175;	Conservative	0;	Mismatches 104;	Indels 4; Gaps
QY	843	TACTACCTCCATATTAAAGTATGACGCCGTTCACATTTTTCTCCAACGGTTGACCATTGC	902		
Db	3080	TACTCTCTCGTTTTTAATAGATGACGCCGTTGACITTTTTCTCCCANTGTTGAOCCATTTC	3021		
QY	903	GTCWTATATAAAAAAATATGCTPATCATTATATTTTATTATATGACITCGTCAATCA	962		
Db	3020	GTCWTATATAAAAATTTATGTAATATGATAAGATATAATCACACTTAAAGTAGCTATGGA	2961		
QY	963	AATATTTTTTAAAGCATGACATAACAATTTTCATATTTTGCAAA-----AAAAAAAACGGAT	1018		
Db	2960	GTGATAAACAAATTCATAACAACTAAATATAATATAATATTTTAAATAAGACGAAAT	2901		
QY	1019	GGTCAAATGTTAGTCGAAAGTCAATGGGTGTCATACATATAAATACGAGGGATTAATG	1078		
Db	2900	GSTCAACATGTGAGAAAAAGTCAACGGCGTCTCTATATAAAAAACAGAGTAGTATATTC	2841		
QY	1079	GGATAGGTACAGTAGCAGTCAATGATGAAGTAGTAATGTGTT	1121		
Db	2840	TTATATACAACTCAATCTACTATCTATGCTTAAGCAAAAAATATTTTT	2798		

RESULT 12	
AAAD38803	
ID	AAAD38803 standard; DNA; 91552 BP.
XX	
AAAD38803;	
XX	
23-SEP-2002	(first entry)
DT	
XX	
DE	BAC clone K6P36 from rice variety CO39.
XX	
KW	Rice; chromosome 11; Indica rice cultivar; CO39; avirulence gene;
KW	AVR1-CO39; rice blast pathogen; transgenic plant; plant breeding;
KW	resistance; agricultural; horticultural; plant protectant; ds.
XX	
OS	Oryza sativa.
XX	
WO200234927-A2.	
XX	
02-MAY-2002.	
XX	
19-OCT-2001; 2001WO-US46331.	
XX	
20-OCT-2000; 2000US-242313P.	
PR	09-JUL-2001; 2001US-303897P.
XX	
PA	(WISC ) WISCONSIN ALUMNI RES FOUND.
XX	(USDA ) US DEPT OF AGRICULTURE.
PA	
XX	
PI	Leong SA, Farman ML, Chauhan RS, Durfee TJ;
XX	
DR	WPI; 2002-471442/50.
XX	
PT	New P1CO39(t) polynucleotides from chromosome 11 of Indica rice
XX	cultivar CO39 useful for conferring or improving resistance of plants
PT	to strains of Magnaporthe grisea or other plant pathogens with
PT	avirulence gene AVR1-CO39 -
XX	
PS	Example 2; Page 68-110; 175pp; English.

4A CC The invention relates to a polynucleotide isolated from chromosome 11 of  
CC Indica rice cultivar CO39, flanked by marker R2316 and RG1094 comprising  
CC one or more genes that confer resistance to strains of Magnaporthe grisea  
CC having avirulence gene AVR1-CO39. The pi-CO39(t) polynucleotides are  
CC useful for conferring or improving resistance of plants to strains of the  
CC rice blast pathogen, Magnaporthe grisea and other pathogens. The  
CC pi-CO39(t) nucleic acids may be used as probes to detect the presence of  
CC and/or expression of pi-CO39(t) genes, and to produce large quantities of  
CC

CC pure Pi-CO39(t) proteins. Purified gene products of Pi-CO39(t) are useful  
 CC in producing polyclonal or monoclonal antibodies useful as sensitive  
 CC detection reagents for the presence and accumulation of Pi-CO39(t)  
 CC polypeptides. The transgenic plants are useful for plant breeding or  
 CC directly in agricultural or horticultural applications. The present  
 CC sequence is BAC clone K6P36 from rice variety CO39.

XX Sequence 91552 BP; 26676 A; 19760 C; 17854 G; 27262 T; 0 other;

Query Match 4.5%; Score 98.2; DB 24; Length 91552;  
 Best local similarity 65.6%; Pred. No. 9.8e-11;  
 Matches 162; Conservative 0; Mismatches 78; Indels 7; Gaps 1;

QY 836 TACTATTACTACCTCCATATTTAAGTATGACGCCGCTTCACTTTTGCCAAAGCTTG 895

DB 33565 TAGTAAGTATCCCTCCGATTTAAGTATGACGCCGCTTCACTTTTGCCAAAGCTTG 33624

QY 896 ACCATTGCTCTTATTAATAAAATATGATATCTATTTATTTATTTATGACTTGATTC 955

DB 33625 ACCATTGCTTTTATTCATAATTTTGCAATATATAAATATTTATGCTTTAAG 33684

QY 956 GTCATCAATATTTTAAAGCATGACATTAACATTTTCATATTTGCAAAAAA-----AA 1008

DB 33685 AATATTGATGACGAATCAAGTCATATAATAAATAAATGATATTAACATAAATTTTGA 33744

QY 1009 AAAACGAAATGTCGAATGTTAGTGAAGTCAATGGTGTGATACATTAATAACGGAG 1068

DB 33745 TAGACGAATGTCGAATGTTAGTGAAGTCAATGGTGTGATACATTAATAACGGAG 33804

QY 1069 GGATTAAT 1075

DB 33805 GGATTAAT 33811

# RESULT 13

ID AAC48763 standard; DNA; 694 BP.

XX AAC48763;

AC AAC48763;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 58679.

KW Hybridisation assay; genetic mapping; gene expression control;  
 KW protein identification; signal transduction pathway;  
 KW metabolic pathway; promoter; termination sequence; ss.

OS Arabidopsis thaliana.

PN EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 21-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 17-JUN-1999; 99US-0139494.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.



```

PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 4.3%; Score 95.4; DB 21; Length 695;
Best Local Similarity 67.2%; Pred. No. 4.6e-11;
Matches 135; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 1693 TGTGATGACGGTGGTCGCGACCAAGGTGTCGAGATGCCGCGGTGTCGGTGGAGGAGG 1752
DB 240 TGATCAAGAGATGTCGCTACCAAGACTTTCGAGATGCCACCATGACTGTCGCTGAGG 299

QY 1753 CGATGGACGACGTAGTGAATGTGGACCAACTTCTACGCTTCAGAGACGAGAGACCG 1812
DB 300 CAGTCGACGACGTGGAAGTACGTCAGTCACGACTTCTATGCTTCCAAAATGAAGAACTG 359

QY 1813 GGGAGATGAAGTCTCTGACAGAGAGAGAGGAGGCTTCGCTCTCATGCTCCCGCAAGG 1872
DB 360 GTGAGATGAACATAGTGTACAGAGAGAGAGAGGAGGTTACGCTCTCATATCCCAAGA 419

QY 1873 GAGACGGTCAFTCCCAAGG 1893
DB 420 AAGACGGGAAGGCGGAGAGG 440

```

## RESULT 15

AAT93780  
ID AAT93780 standard; DNA; 1591 BP.

XX AC AAT93780;

DT 13-FEB-1998 (first entry)

XX DE Oryza sativa pathogen inducible regulatory element Pr-10c cDNA.

XX KW Pathogen inducible regulatory element; MAG-7; Magnaporthe grisea;  
KW promoter region; gene induction; pathogen-inducible regulatory element;  
KW Co39 rice seedling genomic library; disease resistance; rice; ds.

XX OS Oryza sativa.

XX PN US5677175-A.

XX PD 14-OCT-1997.

XX PF 11-OCT-1996; 96US-0005362.

XX PR 13-OCT-1995; 95US-0005362.

XX PA (PURD ) PURDUE RES FOUND.

XX PI Hodges TK, McGee JD;

XX DR WPI; 1997-511872/47.

XX

PT Plant expression vector with pathogen-inducible regulatory element -  
used for increasing disease resistance of plants

XX Claim 1; Columns 11-14; 8pp; English.

XX The present sequence, Pr-10c, is a pathogen inducible regulatory  
CC element. It was identified by screening a Co39 rice seedling genomic  
CC library with a cDNA clone, MAG-7, of a gene whose transcripts appeared in  
CC Co39 rice seedlings 72 hours after inoculation with Magnaporthe grisea.  
CC Induced expression of the Pr-10c gene was not detected after infection  
CC with M. grisea throughout a 144 hour time period. The Pr-10c sequence is  
CC used to make a novel plant expression vector comprising a  
CC pathogen-inducible regulatory element and a core promoter, where the  
CC regulatory element is selected from three rice genomic DNA sequences  
CC AAT93778-80. The vectors are used for increasing the disease resistance  
CC of plants.

XX SQ Sequence 1591 BP; 464 A; 289 C; 373 G; 465 T; 0 other;

Query Match 4.2%; Score 91.4; DB 18; Length 1591;

Best Local Similarity 66.9%; Pred. No. 5e-10;

Matches 164; Conservative 0; Mismatches 71; Indels 10; Gaps 2;

QY 841 ATTACTACCTCCATATTTTAATGATGACGCCGCTCACTTTTGT---CCACGTTTGAC 897

DB 723 AATACTCCTCCGCTTCAAAATGTTGACCCGCTGATTTTTTTTGTATATGTTGAT 782

QY 898 CATTGCTCTTATTAATAAAAAAATAATGATCTATCTATTATTATTATTATGACUTGATTCGT 957

DB 783 CATTGCTCTTATTAATAAAAAAATAATGATCTATCTATTATTATTATTATGACUTGATTCGT 842

QY 958 CATCAAAATATTTTATGACATGACATAAACAATTTTCATATTTTGCAGAAA-----AAAAA 1010

DB 843 TCTTAATAATATCTTTCATGTGACATATAGTTTACATATTTTACATAATTTTTCAGAAATTTTGAATA 902

QY 1011 AAACCAATGTCCTCAATGTTAGTCGAAAACCTCAATGCTCATACATTAATAAATACGAGGG 1070

DB 903 AGACCAAGCGTCAAAACATGTGCTAAAAAATTAAGGTCATATATTTTGAACCGAGGG 962

QY 1071 ATTAT 1075

DB 963 AGTAT 967

Search completed: March 26, 2003, 05:14:07

Job time : 777.159 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2003, 19:50:00 ; Search time 2242.74 Seconds  
(without alignments)  
15894.084 Million cell updates/sec

Title: US-09-702-134-7212\_COPY\_10200\_12400

Perfect score: 2201  
Sequence: 1 cgtcctgtgtttgtgcata.....tagcttgcgtttgtcgtga 2201

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

#### Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estnu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_hic:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_hic:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: gb\_gss:\*

18: em\_gss\_hum:\*

19: em\_gss\_inv:\*

20: em\_gss\_pln:\*

21: em\_gss\_vit:\*

22: em\_gss\_fun:\*

23: em\_gss\_mam:\*

24: em\_gss\_mus:\*

25: em\_gss\_other:\*

26: em\_gss\_pro:\*

27: em\_gss\_rod:\*

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	582.6	26.5	714	17	AQ051086 nbx00005a
c 2	559.6	25.4	653	13	BI305800 NL1_M15
3	529	24.0	638	14	BQ908458 T008E03 O
4	491.4	22.3	597	9	AU094078 AU094078
5	449	20.4	570	9	AU056564 AU056564
6	445.6	20.2	575	13	BI810180 J001B06 O

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

7	443	20.1	522	13	BM420118
8	432	19.6	534	13	BM420133
9	429.4	19.5	525	14	BQ908495
10	418.6	19.0	460	10	BE230575
11	416.8	18.9	627	13	BM037750
12	414.4	18.8	443	14	BQ908635
13	414.0	18.6	516	13	BI811299
14	398.8	18.1	539	13	BM038608
15	397.4	18.1	531	14	BQ906047
16	395	17.9	581	13	BI810545
17	389.4	17.7	576	13	BI811225
18	387	17.6	509	13	BM038731
19	382.4	17.4	515	13	BM421616
20	378.6	17.2	483	13	BI810402
21	377.6	17.2	598	14	BQ907480
22	368	16.7	437	14	C72803
23	362	16.4	399	13	BI806502
24	361.2	16.4	603	10	BE229715
25	360.8	16.4	404	14	BC280726
26	352.6	16.0	557	13	BM038147
27	341.6	15.5	514	13	BM039048
28	341	15.5	521	13	BM420069
29	330.4	15.0	540	9	AU056563
30	328.4	14.9	549	14	BQ908614
31	313.4	14.2	490	13	BM419787
32	313	14.2	552	13	BI811322
33	308	14.0	325	9	AU056562
34	301.6	13.7	604	13	BM038823
35	280.2	12.7	397	14	BQ908902
36	278.8	12.7	540	13	BI807063
37	253.6	11.5	466	14	BQ908540
38	248	11.3	513	13	BI810148
39	236.2	10.7	752	12	BF265998
40	233.6	10.6	858	12	BF262651
41	231.8	10.5	568	13	BJ468880
42	230.8	10.5	595	13	BI679710
43	228.2	10.4	757	12	BF627121
44	216.6	9.8	852	12	BF629402
45	214.8	9.8	380	13	BM421344

#### ALIGNMENTS

RESULT 1  
LOCUS AQ051086 714 bp DNA linear GSS 24-MAR-1999  
DEFINITION nbx0005ac11r CUGI Rice BAC Library Oryza sativa genomic clone  
ACCESSION nbx0005E21r, DNA sequence.  
VERSION AQ051086  
KEYWORDS AQ051086.2 GI:4501804  
SOURCE GSS.  
ORGANISM Oryza sativa.  
REFERENCE 1 (bases 1 to 714)  
AUTHORS Wang, R. A. and Dean, R. A.  
TITLE A BAC End Sequencing Framework to Sequence the Rice Genome  
JOURNAL Unpublished (1998)  
COMMENT On Mar 23, 1999 this sequence version replaced gi:3325391.  
Contact: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu  
Seq primer: GGAAACAGCTATGACCATG  
Class: BAC ends  
High quality sequence start: 31  
High quality sequence stop: 250.

AQ051086 714 bp DNA linear GSS 24-MAR-1999  
nbx0005ac11r CUGI Rice BAC Library Oryza sativa genomic clone  
nbx0005E21r, DNA sequence.  
AQ051086  
AQ051086.2 GI:4501804  
GSS.  
Oryza sativa.  
Oryza sativa.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.

1 (bases 1 to 714)  
Wang, R. A. and Dean, R. A.  
A BAC End Sequencing Framework to Sequence the Rice Genome  
Unpublished (1998)  
On Mar 23, 1999 this sequence version replaced gi:3325391.  
Contact: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu  
Seq primer: GGAAACAGCTATGACCATG  
Class: BAC ends  
High quality sequence start: 31  
High quality sequence stop: 250.

## FEATURES

Source

Location/Qualifiers

1. 714

/organism="Oryza sativa"

/strain="Japonica"

/cultivar="Nipponbare"

/db\_xref="taxon:4530"

/clone="mbx0005E2ir"

/clone\_lib="CUGR Rice BAC Library"

/tissue\_type="Leaf"

/lab\_host="E. coli DH10B"

/note="Vector: pBelobAC11; Site.1: HindIII; Site.2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."

BASE COUNT 174 a 207 c 137 g 196 t

ORIGIN

Query Match 26.5%; Score 582.6; DB 17; Length 714;

Best Local Similarity 93.0%; Pred. No. 7.3e-109;

Matches 654; Conservative 0; Mismatches 44; Indels 5; Gaps 4;

QY 67 CTCCTCTCTCTGTATTAGACAGCAACAGTAGCTAGCGAGCAGACTGCGAGA 126

Db 1 CTCCTCTCTCTGTATTAGACAGCAACAGTAGCTAGCGAGCAGACTGCGAGA 60

QY 127 GCAGAGCATGTTCTTCTCTAGCTATTAGTGTGGGGAATAATGGCTCCGCCACCA 186

Db 61 GCAGAGCATGTTCTTCTCTAGCTATTAGTGTGGGGAATAATGGCTCCGCCACCA 120

QY 187 CGCAGCAATGCGCCTCGCGCGCCCTCTTCCCATCACCATCACCATCATCATCATCAAGC 246

Db 121 CGCAGCAATGCGCCTCGCGCGCCCTCTTCCCATCACCATCACCATCATCATCATCAAGC 180

QY 247 AGCTGAGCTCCCTCGCTCTCTCGTCTCGTCCCTCCCTCCCTCGCTCCCTCGCTCCCT 306

Db 181 AGCTGAGCTCCCTCGCTCTCTCGTCTCGTCCCTCCCTCCCTCGCTCCCTCGCTCCCT 240

QY 307 GCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 366

Db 241 GCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300

QY 367 GCGCGTGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 425

Db 301 GCGCGTGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360

QY 426 GCGCGCGCAACGTCAGAGTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 485

Db 361 GCGCGCGCGCAACGTCAGAGTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420

QY 486 TTTCCT 545

Db 421 TTTCCT 478

QY 546 ATGATTAATTTCTAGATTCAT 605

Db 479 ATGATTAATTTCTAGATTCAT 538

QY 606 TAACCTGAAACAGAGGGAGTATATATAAATTCGGTAATGGAACTAGATATTTCTTTTAT 665

Db 539 TAACCGGAACAGAGGGAGTATATCTAATTCGTAAATGGAACACAGATTCGACTCAAT 598

QY 666 AATTGAGCCTTATCAATAGTACTAGTAATTCACAGAACTAGAT-TCTTCTTTATTGGT 724

Db 599 AATTGAGCCTTATCAATAGTACTAGTAATTCACAGAACTAGATTCGACTCAATGGC 658

QY 725 TATATT-AGTAGTAATTTGGTCTCCTACCTAGTAATTCATTTTATATAC 766

Db 659 TATATTGAGTGGGAATCGCTGACCTTAGACATCTATCTCTAC 701

RESULT 2

BI305800/c 653 bp mRNA linear EST 20-JUL-2001

LOCUS NL\_1\_M15 Drought stress (leaf) Oryza sativa cDNA clone NL\_1\_M15 3', mRNA sequence.

DEFINITION

ACCESSION BI305800

VERSION BI305800.1 GI:14981109

KEYWORDS EST.

SOURCE Oryza sativa.

ORGANISM Oryza sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE 1 (bases 1 to 653)

AUTHORS Reddy A.R., Ramakrishna, W., Chandrasekhar, A., Nagabhushan, I., Ravindrababu, P. and Bennetzen, J.L.

TITLE Novel EST enrichment with normalized cDNA libraries from drought stressed rice (Oryza sativa L.cv Nagina 22)

JOURNAL Unpublished (2001)

COMMENT Contact: Reddy AR

Department of Plant Sciences, School of Life Sciences

University of Hyderabad

P.O. Central University, Hyderabad-500 046, A.P., India

Tel: 0091-40-3010265

Fax: 0091-40-3010145

Email: arjuls@uohyd.ernet.in

Insert Length: 653 Std Error: 0.00

Plate: 1 row: M column: 15

Seq primer: GTAAACGAGCGGCAGTG.

Location/Qualifiers

1. 653

/organism="Oryza sativa"

/cultivar="Nagina 22 (indica sub sp)"

/db\_xref="taxon:4530"

/clone="NL\_1\_M15"

/clone\_lib="Drought stress (leaf)"

/tissue\_type="Entire leaf tissue"

/dev\_stage="35 day-old seedlings"

/note="Organ: Leaf; Vector: p713Pac; ESTs from normalized leaf cDNA Library from drought stressed seedlings"

BASE COUNT 144 a 159 c 187 g 162 t

ORIGIN

Query Match 25.4%; Score 559.6; DB 13; Length 653;

Best Local Similarity 88.9%; Pred. No. 3.7e-104;

Matches 648; Conservative 0; Mismatches 5; Indels 76; Gaps 1;

QY 1413 AGCCGAGCATGCTCGCTCCCTCCATCATCAAGAGAACTCAGGAAGATCAAGGAGA 1472

Db 653 AGCCGAGCATGCTCGCTCCCTCCATCATCAAGAGAACTCAGGAAGATCAAGGAGA 594

QY 1473 AGGAGACGAGCTCGGCACTCAAGGGGACCAAGCCCGCTCCGACTGGCGGCCAT 1532

Db 593 AGGAGACGAGCTCGGCACTCAAGGGGACCAAGCCCGCTCCGACTGGCGGCCAT 534

QY 1533 CATTCGTCGACACAGAGCAGCAGCCCGCCAGCCCACTCAAGGATCTGGAGAGGCGG 1592

Db 533 CATTCGTCGACACAGAGCAGCAGCCCGCCAGCCCACTCAAGGATCTGGAGAGGCGG 474

QY 1593 TCGGCGCGGAGGAGGAGCAGCGTGTCTCAACAGGTCTGCAGTCTCATGATCAT 1652



TITLE Rice cDNA from panicle at flowering stage (2000)  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Takuji Sasaki  
 National Institute of Agrobiological Resources  
 Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki  
 305-8602, Japan  
 Tel: 81-298-38-7441  
 Fax: 81-298-38-7468  
 Email: tsasaki@abr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/  
 PROJECT = "RGP".

FEATURES  
 source  
 1. .597  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="E2287"  
 /clone\_lib="Rice panicle at flowering stage"  
 /dev\_stage="flowering stage"  
 /note="Organ: panicle; Rice cDNA from panicle at flowering stage"

BASE COUNT 156 a 159 c 138 g 140 t 4 others  
 ORIGIN

Query Match 22.3%; Score 491.4; DB 9; Length 597;  
 Best Local Similarity 87.7%; Pred. No. 3.3e-90;  
 Matches 578; Conservative 0; Mismatches 5; Indels 76; Gaps 1;

QY 1494 TCAAGGGGACCAAGCCCGCTCCGACTGCGCCGCATCATCTCGACCAACAGGAGC 1553  
 Db 1 TCAAGGGGACCAAGCCCGCTCCGACTGCGCCGCATCATCTCGACCAAGGAGC 60  
 QY 1554 ACAGCCCGAGCCCGAGCTCAAGATCTGGAGAGCGCGCTCGGCGGAGGAGGAGCA 1613  
 Db 61 ACAGCCCGAGCCCGAGCTCAAGATCTGGAGAGCGCGCTCGGCGGAGGAGGAGCA 120  
 QY 1614 CGGTGCTCACCAGCTCTGCACTGTCGATGATCATGATATTAATCAATCATCA 1673  
 Db 121 CGGTGCTCACC----- 132  
 QY 1674 TGAACCAAAATTAATCAACTGTGAATGCAAGTGGTGGCCACCAAGGTGTCGAGATGCGC 1733  
 Db 133 -----AGTGGTGGCCACCAAGGTGTCGAGATGCGC 164  
 QY 1734 CCCTGTCCGTGGAGGAGGAGGAGGAGCTAGTGAATGTGGACCAACACTTCTACGCC 1793  
 Db 165 CCCTGTCCGTGGAGGAGGAGGAGGAGCTAGTGAATGTGGACCAACACTTCTACGCC 224  
 QY 1794 TTCAGAGACGAGAACCGGGAGATGAAGTCCCTGTACAGAGGAGGAGGAGGCTTC 1853  
 Db 225 TTCAGAGACGAGAACCGGGAGATGAAGTCCCTGTACAGAGGAGGAGGAGGCTTC 284  
 QY 1854 GTCTCATCGTCCCAAGGAGGAGGAGTCTCCACAGAGGAGACCATCCCAACTCTGAC 1913  
 Db 285 GTCTCATCGTCCCAAGGAGGAGGAGTCTCCACAGAGGAGACCATCCCAACTCTGAC 344  
 QY 1914 CACACACCCCTCCCTGCTAGCTAGCTCTACACCTCTGATCACCCTCTCTCTTTG 1973  
 Db 345 CACACACCCCTGCTGCTAGCTCTACACCTCTGATCACCCTCTCTCTTTG 404  
 QY 1974 CTAAATATTTCTATGCTTGCACAGAGTCTCTAAATATTACTAGTATTAATTAATCAAG 2033  
 Db 405 CTAAATATTTCTATGCTTGCACAGAGTCTCTAAATATTACTAGTATTAATTAATCAAG 464  
 QY 2034 TGATGAGAGGCTCTTGAATTTCTCCGTGTACATGATCGATCGACTCTTGACTACT 2093  
 Db 465 TGATGAGAGGCTCTTGAATTTCTCCGTGTACATGATCGATCGACTCTTGACTACT 524  
 QY 2094 CTGTTAAACCACTGTTTTCGATATTTATTAATTAATTAATTAATTAATTAATTAATTA 2152  
 Db 525 CTGTTAAACCACTGTTTTCGATATTTATTAATTAATTAATTAATTAATTAATTAATTA 583

RESULT 5

AU056564  
 LOCUS 570 bp mRNA linear EST 01-APR-2002  
 DEFINITION AU056564 Oryza sativa mature leaf Nipponbare Oryza sativa (japonica cultivar-group) cDNA clone S20686\_22, mRNA sequence.  
 ACCESSION AU056564  
 VERSION AU056564.1 GI:4715447  
 KEYWORDS EST.  
 SOURCE Oryza sativa (japonica cultivar-group).  
 ORGANISM Oryza sativa (japonica cultivar-group)  
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.  
 REFERENCE 1 (bases 1 to 570)  
 AUTHORS Yamamoto, K. and Sasaki, T.  
 TITLE Rice cDNA from mature leaf  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Takuji Sasaki  
 National Institute of Agrobiological Resources  
 Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki  
 305-8602, Japan  
 Tel: 81-298-38-7441  
 Fax: 81-298-38-7468  
 Email: tsasaki@abr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/  
 PROJECT = "RGP".

FEATURES  
 Location/Qualifiers  
 source  
 1. .570  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="S20686\_22"  
 /clone\_lib="Oryza sativa mature leaf Nipponbare"  
 /tissue\_type="mature leaf"  
 BASE COUNT 148 a 153 c 125 g 137 t 7 others  
 ORIGIN

Query Match 20.4%; Score 449; DB 9; Length 570;  
 Best Local Similarity 98.7%; Pred. No. 1.6e-81;  
 Matches 449; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1702 AGTGTGTCGGACCAAGGTGTTCAGATGCGCGCTGTCGGTGGAGGAGCGATGGAGC 1761  
 Db 100 AGTGTGTCGGACCAAGGTGTTCAGATGCGCGCTGTCGGTGGAGGAGCGATGGAGC 159  
 QY 1762 AGTGTGTCGGACCAAGGTGTTCAGATGCGCGCTGTCGGTGGAGGAGCGATGGAGC 1821  
 Db 160 AGTGTGTCGGACCAAGGTGTTCAGATGCGCGCTGTCGGTGGAGGAGCGATGGAGC 219  
 QY 1822 ACCTCTGTACAGAGGAGGAGGAGGCTTCGGTCTCATGTCGCCAAGGAGGAGGCTC 1881  
 Db 220 ACCTCTGTACAGAGGAGGAGGAGGCTTCGGTCTCATGTCGCCAAGGAGGAGGCTC 279  
 QY 1882 ATCTCACAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1941  
 Db 280 ATCTCACAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 339  
 QY 1942 CTCTACACCTCTCTGATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2001  
 Db 340 CTCTACACCTCTCTGATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 399  
 QY 2002 CTCTAAATATTTACTAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2061  
 Db 400 CTCTAAATATTTACTAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 459  
 QY 2062 CGTGTACATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCG 2121  
 Db 460 CGTGTACATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCG 519  
 QY 2122 TATATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2156  
 Db 520 TATATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 554

RESULT 6

QY	2037	TCGAGAGGCTCTTGATTTTCCTCGGTGCATCATCGATCGACTCTTGACTACTCTG	2099
Db	479	TCGAGAGGCTCTTGATTTTCCTCGGTGCATCATCGATCGACTCTTGACTACTCTG	538
QY	2097	TTTAACCACTCTTTTCGATATATTTATATTATTA	2132
Db	539	TTTAACCACTCTTTTCGATATATTTATATTATTA	574
RESULT 7			
BM420118		522 bp mRNA	Linear EST 28-JAN-2000
LOCUS			
DEFINITION	BM420118	R023H11 Oryza sativa mature leaf library induced by M.g:risea Or	
ACCESSION	BM420118	sativa cDNA clone R023H11, mRNA sequence.	
VERSION	BM420118.1	GI:18386920	
KEYWORDS	EST.		
SOURCE	Oryza sativa		
ORGANISM	Oryza sativa		
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
	1..522		
	/organism="Oryza sativa"		
	/db_xref="taxon:4530"		
	/clone="R023H11"		
	/clone_lib="Oryza sativa mature leaf library induced b		
	M.g:risea"		
	/tissue_type="leaf"		
	/dev_stage="Mature stage"		
	/note="vector: pSport2"		
BASE COUNT	138 a 136 c 122 g 126 t		
ORIGIN			
Query Match	20.18;	Score 443;	DB 13; Length 522;
Best Local Similarity	100.0%;	Pred. No. 2.6e-80;	
Matches 443;	Conservative 0;	Mismatches 0;	Indels 0; Gaps
QY	1702	AGGTGGTGCACCAAGTGTTTGAGATCCCGCGCTCGTGGTGAGGAGCGATGGAG	1761
Db	68	AGGTGGTGCACCAAGTGTTTGAGATCCCGCGCTCGTGGTGAGGAGCGATGGAG	127
QY	1762	AGCTAGTGAATGTGGACCACTTCCTACGCCCTTCAGACAGAGAGACCGGGGAGATGA	1821
Db	128	AGCTAGTGAATGTGGACCACTTCCTACGCCCTTCAGACAGAGAGACCGGGGAGATGA	187
QY	1822	ACGTCCTGTACAGAGGAGGAGGAGGCGTTCGGTCTCATCGTCCCAAGGAGACGGTC	1881
Db	188	ACGTCCTGTACAGAGGAGGAGGAGGCGTTCGGTCTCATCGTCCCAAGGAGACGGTC	247
QY	1882	ATCTCCACAGGAGACCATCCCAACTCTTGACCCACCAACCCCTCCCTTGCTGCTAGC	1941
Db	248	ATCTCCACAGGAGACCATCCCAACTCTTGACCCACCAACCCCTCCCTTGCTGCTAGC	307
QY	1942	CTCTACACCTCTTGATCACCACCTCTTATTCCTTAATTTCTATGCTGCTGCCACAGT	2001
Db	308	CTCTACACCTCTTGATCACCACCTCTTATTCCTTAATTTCTATGCTGCTGCCACAGT	367

QY	2002	CTCTAAATATTACTAGTATTATTAATTAATCCAAAGTGATGAGAGAGGTCTTGATTTGCTCTC	2061
Db	368	CTCTAAATATTACTAGTATTATTAATTAATCCAAAGTGATGAGAGAGGTCTTGATTTGCTCTC	427
QY	2062	CGTGTACATCGATCGATCGACTCTTGACTACTCTGTTTAACCAACTGTTTCGATATATT	2121
Db	428	CGTGTACATCGATCGATCGACTCTTGACTACTCTGTTTAACCAACTGTTTCGATATATT	487
QY	2122	TATATTATTAAATGATTTTCCAA	2144
Db	488	TATATTATTAAATGATTTTCCAA	510
RESULT	8		
LOCUS	BM420133		
DEFINITION	R024B06 Oryza sativa mature leaf library induced by M.grisea Oryza sativa cDNA clone R024B06, mRNA sequence.		
ACCESSION	BM420133		
VERSION	BM420133.1	GI:16386935	
KEYWORDS	EST.		
SOURCE	Oryza sativa.		
ORGANISM	Oryza sativa		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.		
AUTHORS	1 (bases 1 to 534)		
TITLE	Dong,H.T., Li,D.B., Zhuang,X.F., Dai,C.G., Sun,L.X., Pei,Y.X., Wu,H.F., Jiang,Y.X., Yu,F.C., Gao,Q.K. and Jau,Y.C.		
JOURNAL	A gene expression screen in Oryza sativa		
COMMENT	Unpublished (2001) Contact: Dong HT Laboratory of Functional Genetics Bio-Technology Institute of Zhejiang University Kaixuan Road 368#, Hangzhou, Zhejiang, P.R.China Tel: 0086-571-86892051 Fax: 0086-571-86961525 Email: htdong@zjuem.zju.edu.cn Seq primer: M3 forward primer. Location/Qualifiers 1..534 /organism="Oryza sativa" /db_xref="taxon:4530" /clone="R024B06" /clone_lib="Oryza sativa mature leaf library induced by M.grisea" /tissue_type="leaf" /dev_stage="Mature stage" /note="Vector: pSport2"		
BASE COUNT	142 a 139 c 127 g 126 t		
ORIGIN			
Query Match	19.6%;	Score 432;	DB 13; Length 534;
Best Local Similarity	99.8%;	Pred. No. 4.7e-78;	
Matches 443;	Conservative 0;	Mismatches 0;	Indels 1; Gaps 1;
QY	1702	AGTGTGGTGGCCACCAAGGTGTCGAGATGCCCGCGTGTGCGTGAGAGAGCGATGGAGC	1761
Db	78	AGTGTGGTGGCCACCAAGGTGTCGAGATGCCCGCGTGTGCGTGAGAGAGCGATGGAGC	137
QY	1762	AGCTAGTAGTGTGGACCAACTTCTACGCCCTTCAGACACGAGAAGACCGGGGATGTA	1821
Db	138	AGCTAGTAGTGTGGACCAACTTCTACGCCCTTCAGACACGAGAAGACCGGGGATGTA	197
QY	1822	AGTCTCTTACAGAGAAGAGAGAGGCGTTTCGGTGTCTATTCGTCGCCCAAGGGAGCGGTC	1881
Db	198	AGTCTCTTACAGAGAAGAGAGAGGCGTTTCGGTGTCTATTCGTCGCCCAAGGGAGCGGTC	257
QY	1882	ATCTCCACAGAGACCATCCCACTCTGACCAACCAACCCCTCCCTTCCTGCTAGC	1941
Db	258	ATCTCCACAGAGACCATCCCACTCTGACCAACCAACCCCTCCCTTCCTGCTAGC	317
QY	1942	CTCTACACCTCCTGCATACCACTCTCTATTTCGTTAAATTATTTCTATGTCGCCCAAGGT	2001

D	b	318	CCTCTACACTCCTCGCATACCATCCTCAITTTGGTAAATTATTTCTTAAGTCCTGCCAAGACT	377
Q	y	2002	CCTCAAATATTAC-TAGTATTAAATTAATCAAGTGATGAGAGAGGTCTTTGATTGTCT	2060
D	b	378	CCTCAAAATATTACGTAGTAGTATTAAATCAAGTGAAGAGAGGTCTTTGATTGTCT	437
Q	y	2061	CCGTGCATCATGCATGCATGCATCTTGACTACTCTGTTAAACCAACTGTTTTCGATATAT	2120
D	b	438	CCGTGCATCATGCATGCATGCATCTTGACTACTCTGTTAAACCAACTGTTTTCGATATAT	497
Q	y	2121	TTATATTATTAAATGATTTTCCAA	2144
D	b	498	TTATATTATTAAATGATTTTCCAA	521
RESULT 9				
BQ908495				
LOCUS				
DEFINITION BQ908495 525 bp mRNA linear EST 19-AUG-2002				
T008H05 Oryza sativa mature leaf library induced by M.grisea Oryza				
sativa cDNA clone T008H05, mRNA sequence..				
ACCESSION BQ908495				
VERSION BQ908495.1 GI:22307259				
KEYWORDS EST.				
SOURCE Oryza sativa.				
ORGANISM Oryza sativa				
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;				
Ehrnartoideae; Oryzaleae; Oryza.				
REFERENCE 1 (bases 1 to 525)				
AUTHORS Dong,H.T., Li,D.B., Zhuang,X.F., Dai,C.G., Sun,L.X., Pei,Y.X., Wu				
,H.F., Jiang,Y.X., Yu,F.C., Gao,Q.K. and Lou,Y.C.				
TITLE A Gene Expression Screen in Oryza sativa				
JOURNAL Unpublished (2001)				
COMMENT Contact: Dong HT				
Laboratory of Functional Genetics				
Bio-technology Institute of Zhejiang University				
Kaixuan Road 368#, Hangzhou, Zhejiang, P.R.China				
Tel: 0086-571-86892051				
Fax: 0086-571-86361525				
Email: htdong@zjuem.zju.edu.cn				
Seq primer: M13 forward primer.				
FEATURES				
source				
i . 525				
/organism="Oryza sativa"				
/db_xref="taxon:4530"				
/clone_lib="T008H05"				
M.grisea"				
/tissue_type="leaf"				
/dev_stage="Mature stage"				
/note="Vector: pSport2"				
BASE COUNT 142 a 134 c 120 g 128 t 1 others				
ORIGIN				
Query Match 19.5%; Score 429.4; DB 14; Length 525;				
Best Local Similarity 99.3%; Pred. No. 1.6e-77;				
Matches 441; Conservative 0; Mismatches 2; Indels 1; Gaps 1;				
Q	y	1702	AGGTGTCGCACCAAGAGTGTTCGAGATCGCGCGGTGTCGGTGGAGGAGCGATGGAGC	1761
D	b	63	AGNTGTCGCACCAAGAGTGTTCGAGATCGCGCGGTGTCGGTGGAGGAGCGATGGAGC	122
Q	y	1762	AGCTASTG-ANTGTGGACCAACAACCTTACGCCCTTGAGAGAGAGAAGCCGGGAGATG	1820
D	b	123	AGCTASTGTAATGTGGACCAACAACCTTACGCCCTTGAGAGAGAGAAGCCGGGAGATG	182
Q	y	1821	AACGTCCTGTACAGAGGAGAGAGAGGCTTCGGTCTCATGTCCTCCCAAGGAGAGCGGT	1880
D	b	183	AACGTCCTGTACAGAGGAGAGAGAGGCTTCGGTCTCATGTCCTCCCAAGGAGAGCGGT	242
Q	y	1881	CATCTCCCAAGGAGAGCATCCCCAACCTCTGACCAACACACCCCTCCTCTGCTGCTAG	1940



```

Db 243 CACTCCACAGGAGACCATCCCACTCTGACCAACACCCCTCCCTCTGCTGCTAG 302
QY 1941 CTTCTACACCTCTGCAATCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2000
Db 303 CTTCTACACCTCTGCAATCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 362
QY 2001 TCTCTAAATATATCTAGTATATATATATATATATATATATATATATATATAT 2060
Db 363 TCTCTAAATATATCTAGTATATATATATATATATATATATATATATATATAT 422
QY 2061 CCGTGATACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2120
Db 423 CCGTGATACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 482
QY 2121 TTATATATATATATATATATATATATATATATATATATATATATATATAT 2144
Db 483 TTATATATATATATATATATATATATATATATATATATATATATATATAT 506

RESULT 10
BE230575 460 bp mRNA linear EST 07-JUL-2000
LOCUS 99AS79 Rice Seedling Lambda ZAPII cDNA Library Oryza sativa cDNA
DEFINITION clone 99AS79, mRNA sequence.
ACCESSION BE230575
VERSION BE230575.1 GI:8956772
KEYWORDS EST.
SOURCE Oryza sativa.
ORGANISM Oryza sativa
REFERENCE 1 (bases 1 to 460)
AUTHORS Lee,M.C., Shin,Y.C., Lee,T.H., Jeong,S.H., Kim,J.K., Eun,M.Y. and
Nahm,B.H.
TITLE Large-scale Sequencing Analysis of ESTs from Rice Seedling
JOURNAL Unpublished (1999)
COMMENT Contact: Eun M.Y.
Department of Cyto Genetics
National Inst. of Agri. Sci. and Tech, RDA
Suwon, Kyunggido, Korea
Tel: 82 331 290 0301
Fax: 82 331 290 0307
Email: myeun@sun20.asti.re.kr.
FEATURES
source
1..460
/organism="Oryza sativa"
/cultivar="Milyang23"
/db_xref="taxon:4530"
/clone="99AS79"
/clone_lib="Rice seedling Lambda ZAPII cDNA Library"
/dev_stage="5 days after pollination"
/lab_host="E. coli SOLR"
/notes="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
XhoI; Directional cDNA library inserted into lambda ZAPII
vector at 5' end with EcoRI and 3' end with Xho I site"
BASE COUNT 81 a 161 c 111 g 107 t
ORIGIN
Query Match 19.0%; Score 418.6; DB 10; Length 460;
Best Local Similarity 97.8%; Pred. No. 2.6e-75;
Matches 435; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

QY 1 CGTCGTCGTTTGGCATCCATCCAGATTCGTTCCAGATCCAGCAGCCGATCTC 60
Db 11 CGTCGTCGTTTGGCATCCAGATTCGTTCCAGATCCAGCAGCCGATCTC 70
QY 61 AAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
Db 71 AAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 130
QY 121 GCGAGACAGACATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180

```

```

Db 131 GCAGAGCAGAGCATTTGTTCTTCTCTAGTATTAGTGTGGGGAATATGGCTCCG 190
QY 181 CCACCAACCGCAGCAATGGCCCTCGCGCCGCTCTCTCTCTCTCTCTCTCTCTCT 240
Db 191 CCACCAACCGCAGCAATGGCCCTCGCGCCGCTCTCTCTCTCTCTCTCTCTCTCT 250
QY 241 TCAGCAGCTGAGCTCCCTCGCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
Db 251 TCAAGCAGCTGAGCTCCCTCGCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 310
QY 301 TTGGCCGCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 311 TTGGCCGCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 369
QY 361 GCAGCAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 370 GCATCATCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 429
QY 421 TTATGACGGCGCGCAAGCTCAAGGT 445
Db 430 TTATGACGGCGCGCAAGCTCAAGGT 454

RESULT 11
BM037750 627 bp mRNA linear EST 06-NOV-2001
LOCUS S098C08 Stem library from Oryza sativa (3-5 leaf stage) Oryza
DEFINITION sativa cDNA clone S098C08, mRNA sequence.
ACCESSION BM037750
VERSION BM037750.1 GI:16753371
KEYWORDS EST.
SOURCE Oryza sativa.
ORGANISM Oryza sativa
REFERENCE 1 (bases 1 to 627)
AUTHORS Dong,H.T., Li,D.B., Zhuang,X.F., Dai,C.G., Sun,L.X., Pei,Y.X., Wu
,H.F., Jiang,Y.X., Yu,F.C., Gao,Q.K. and Lou,Y.C.
TITLE A Gene Expression Screen in Oryza sativa
JOURNAL Unpublished (2001)
COMMENT Contact: Dong HT
Laboratory of Functional Genetics
Bio-technology Institute of Zhejiang University
Kaixuan Road 268#, Hangzhou, Zhejiang, P.R.China
Tel: 0086-571-86892051
Fax: 0086-571-86961525
Email: htdong@zjuem.zju.edu.cn
Seq primer: M13 forward primer.
FEATURES
source
1..627
/organism="Oryza sativa"
/db_xref="taxon:4530"
/clone="S098C08"
/clone_lib="stem library from Oryza sativa (3-5 leaf stage)"
/tissue_type="Stem"
/dev_stage="3-5 leaf stage"
/notes="Vector: pSport2"
BASE COUNT 161 a 171 c 145 g 150 t
ORIGIN
Query Match 18.9%; Score 416.8; DB 13; Length 627;
Best Local Similarity 99.1%; Pred. No. 6e-75; 2; Indels 2; Gaps 2;
Matches 440; Conservative 0; Mismatches 0;

QY 1702 AGGTGTTGGCCACCAAGTGTTCAGATGCCGCCGCTCTCTGCTGAGAGCGATGAGC 1761
Db 175 AGGTGTTGGCCACCAAGTGTTCAGATGCCGCCGCTCTCTGCTGAGAGCGATGAGC 234
QY 1762 AGCTAGTGAATGTGGACCAACTTCTACCCCTTCAGAGACGAGAGACCGGGAGATGA 1821
Db 235 AGCTAGTGAATGTGGACCAACTTCTACCCCTTCAGAGACGAGAGACCGGGAGATGA 294

```

```

QY 1822 AGTCTCTGTACAGAGGAGGAGGCTTCGGTCTCATCTCCCTCCCAAGGAGAGGCTC 1881
Db 295 AGTCTCTGTACAGAGGAGGAGGCTTCGGTCTCATCTCCCTCCCAAGGAGAGGCTC 354
QY 1882 ATCTCCA-CAGAGGAGGAGGAGGCTTCGGTCTCATCTCCCTCCCAAGGAGAGGCTC 1940
Db 355 ATCTCCAAGGAGGAGGAGGCTTCGGTCTCATCTCCCTCCCAAGGAGAGGCTC 414
QY 1941 CCTCTACACCTCCCTGTCATCATCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 2000
Db 415 CCTCTACACCTCCCTGTCATCATCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 474
QY 2001 TCTCTAATATATCTAGTATTAATTAATCAAGTATCAAGGAGGCTTCGGTCTCCCT 2060
Db 475 TCTCTAATATATCTAGTATTAATTAATCAAGTATCAAGGAGGCTTCGGTCTCCCT 534
QY 2061 CGGTGTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2120
Db 535 CGGTGTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 593
QY 2121 TTATATTAATATGATTTTCCAA 2144
Db 594 TTATATTAATATGATTTTCCAA 617

RESULT 12
BQ908635
LOCUS
DEFINITION
T012B09 Oryza sativa mature leaf library induced by M.grisea Oryza
sativa cDNA clone T012B09, mRNA sequence.
ACCESSION
BQ908635
VERSION
BQ908635.1 GI:22307413
KEYWORDS
EST.
SOURCE
Oryza sativa.
ORGANISM
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 443)
Dong,H.T., Li,D.B., Zhuang,X.F., Dai,C.G., Sun,L.X., Pei,Y.X., Wu
,H.F., Jiang,Y.X., Yu,F.C., Gao,Q.K. and Lou,Y.C.
A Gene Expression Screen in Oryza sativa
Unpublished (2001)
Contact: Dong HT
Laboratory of Functional Genetics
Bio-technology Institute of Zhejiang University
Kaixuan Road 268#, Hangzhou, Zhejiang, P.R.China
Tel: 0086-571-86961525
Fax: 0086-571-86961525
Email: htdong@zjuem.zju.edu.cn
Seq primer: M13 forward primer.
FEATURES
Location/Qualifiers
1..443
/organism="Oryza sativa"
/db_xref="taxon:4530"
/clone_lib="T012B09"
M.grisea
/tissue_type="leaf"
/dev_stage="Mature stage"
/notes="Vector: pSport2"
BASE COUNT 122 a 113 c 93 g 115 t
ORIGIN
Query Match 18.8%; Score 414.4; DB 14; Length 443;
Best Local Similarity 99.5%; Pred. No. 1.9e-74;
Matches 426; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1711 GCACCAAGGTCTTCAGATGCCCGCTCGTGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1770
Db 1 GCACCAAGGTCTTCAGATGCCCGCTCGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 60

```

```

QY 1771 ATGTGGACCAACAATCTTACGGCTTCAGAGACGAGAAAGCGGGAGATGAACGCTCTGT 1830
Db 61 ATGTGGACCAACAATCTTACGGCTTCAGAGACGAGAAAGCGGGAGATGAACGCTCTGT 120
QY 1831 ACA-AGAGAGAGGAGGAGGCTTCGGTCTCATCTCCCTCCCAAGGAGAGGCTCATCTCC 1889
Db 121 ACACAGAGAGGAGGAGGCTTCGGTCTCATCTCCCTCCCAAGGAGAGGCTCATCTCC 180
QY 1890 AAGGAGACATCCCAACTCTGACCAACCAACCAACCCCTCCCTGCTAGCCTCTACAC 1949
Db 181 AAGGAGACATCCCAACTCTGACCAACCAACCAACCCCTCCCTGCTAGCCTCTACAC 240
QY 1950 CTCCTGATCACCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2009
Db 241 CTCCTGATCACCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
QY 2010 ATTACTAGTATTAATTAATCAAGTATCAAGGAGGCTTCGGTCTCCCTGAGTACA 2069
Db 301 ATTACTAGTATTAATTAATCAAGTATCAAGGAGGCTTCGGTCTCCCTGAGTACA 360
QY 2070 TCGATCGATCGATCTGACTCTCTGTTTAAACCAACTGTTTCGATATATATATAT 2129
Db 361 TCGATCGATCGATCTGACTCTCTGTTTAAACCAACTGTTTCGATATATATATAT 420
QY 2130 TAATGATT 2137
Db 421 TAATGATT 428

RESULT 13
B1811299
LOCUS
DEFINITION
B1811299 Oryza sativa mature leaf library induced by M.grisea Oryza
sativa cDNA clone N001C12, mRNA sequence.
ACCESSION
B1811299
VERSION
B1811299.1 GI:15058487
KEYWORDS
EST.
SOURCE
Oryza sativa.
ORGANISM
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 516)
Dong,H.T., Li,D.B., Zhuang,X.F., Dai,C.G., Sun,L.X., Pei,Y.X., Wu
,H.F., Jiang,Y.X., Yu,F.C., Gao,Q.K. and Lou,Y.C.
A Gene Expression Screen in Oryza sativa
Unpublished (2001)
Contact: Dong HT
Laboratory of Functional Genetics
Bio-technology Institute of Zhejiang University
Kaixuan Road 268#, Hangzhou, Zhejiang, P.R.China
Tel: 0086-571-86961525
Fax: 0086-571-86961525
Email: htdong@zjuem.zju.edu.cn
Seq primer: M13 forward primer.
FEATURES
Location/Qualifiers
1..516
/organism="Oryza sativa"
/db_xref="taxon:4530"
/clone_lib="N001C12"
M.grisea
/tissue_type="leaf"
/dev_stage="Mature stage"
/notes="Vector: pSport2"
BASE COUNT 142 a 134 c 114 g 126 t
ORIGIN
Query Match 18.6%; Score 410; DB 13; Length 516;
Best Local Similarity 98.2%; Pred. No. 1.5e-73;
Matches 436; Conservative 0; Mismatches 5; Indels 3; Gaps 2;
QY 1701 CAGGTGTTCGCCACCAAGGTTCGAGATGCCCGCTGTGCTGGAGGAGGAGGAGGAG 1760

```

Query Match 18.1%; Score 398.8; DB 13; Length 539;  
Best Local Similarity 97.5%; Pred. No. 2.9e-71;

BASE COUNT	141 a	140 c	113 g	136 t	1 others
ORIGIN					

```

Query Match      18.1%; Score 397.4; DB 14; Length 531;
Best Local Similarity 99.5%; Pred. No. 5.6e-71;
Matches 409; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1702 AGGTGGTGGCCACCAAGGTGTTTCAGATGCCGCCGCTGTCGGTGGAGAGCGCATGGAGC 1761
Db      |||||||
QY 1762 AGCTAGTGAATGTGGACCACTTCTACGCTTCAGAGACGAGAGACCGGGGAGATGA 1821
Db      |||||||
QY 99 AGCTAGTGAATGTGGACCACTTCTACGCTTCAGAGACGAGAGACCGGGGAGATGA 158
Db      |||||||
QY 1822 ACCTCCTGTACAAAGAGAGAGAGGCTTCGGTCTCATCGTCCCAAGGGAGACGGTC 1881
Db      |||||||
QY 159 ACCTCCTGTACAAAGAGAGAGAGGCTTCGGTCTCATCGTCCCAAGGGAGACGGTC 218
Db      |||||||
QY 1882 ATCTCCACAAGAGAGACATCCCAACTCTGACACACCCACCCCTCCCTGCTGCTAGC 1941
Db      |||||||
QY 219 ATCTCCACAAGAGAGACATCCCAACTCTGACACACCCACCCCTCCCTGCTGCTAGC 278
Db      |||||||
QY 1942 CTCTACACCTCCTGCATCACCCTCCTCATTTGCTAATTATTCTATGCTTGCCCAAGAGT 2001
Db      |||||||
QY 279 CTCTACACCTCCTGCATCACCCTCCTCATTTGCTAATTATTCTATGCTTGCCCAAGAGT 338
Db      |||||||
QY 2002 CTCTAAATATTACTAGTATTAAATTAATCAAGTGAAGAGAGTCTTTGATTTGCTC 2061
Db      |||||||
QY 339 CTCTAAATATTACTAGTATTAAATTAATCAAGTGAAGAGAGTCTTTGATTTGCTC 398
Db      |||||||
QY 2062 CGTGTACATCGATCGATCGACTCTTGACTACTCTGTTTAAACCAACTGTTT 2112
Db      |||||||
QY 399 CGTGTACATCGATCGATCGACTCTTGACTACTCTGTTTAAACCAACTGTTT 449

```

Search completed: March 26, 2003, 07:52:29  
Job time : 2267.4 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2003, 04:38:30 ; Search time 54.4409 Seconds  
(without alignments)  
12398.686 Million cell updates/sec

Title: US-09-702-134-7212\_COPY\_10200\_12400  
Perfect score: 2201  
Sequence: 1 cgtcttgctttgtgcatc.....tagcttgctttgtgtgta 2201

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_NA.\*  
1: /cgn2.6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2.6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2.6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2.6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2.6/ptodata/1/ina/PCrus\_COMB.seq.\*  
6: /cgn2.6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	91.4	4.2	1591	1	US-08-728-956-3
2	80.8	3.7	2317	4	US-09-105-390-3
3	77	3.5	7218	1	US-08-232-463-14
4	74.2	3.4	7218	1	US-08-232-463-14
5	62.2	2.8	636	4	US-08-998-416-1137
6	61.8	2.8	1591	1	US-08-728-956-3
7	59.6	2.7	1138	3	US-08-581-1480-3
8	58.8	2.7	2317	4	US-09-105-390-3
9	58.4	2.7	615	4	US-08-998-416-186
10	56.6	2.6	390	4	US-09-197-649-7
11	56	2.5	837	4	US-08-998-416-288
12	52.8	2.4	51259	3	US-08-781-891-209
13	51.2	2.3	6124	4	US-08-213-419B-3
14	50.8	2.3	1140	3	US-09-023-173-4
15	50.6	2.3	19124	2	US-08-487-826B-13
16	50	2.3	665	2	US-08-883-795A-36
17	49.2	2.2	837	4	US-08-998-416-288
18	49.2	2.2	1182	4	US-09-385-028-19
19	49.2	2.2	1926	4	US-09-249-585A-4
20	49.2	2.2	1931	2	US-09-130-114-2
21	49.2	2.2	11604	4	US-09-385-028-13
22	49.2	2.2	15079	4	US-09-385-028-1
23	49	2.2	711	4	US-08-998-416-786
24	49	2.2	33529	4	US-09-144-085-3
25	48.4	2.2	19124	2	US-08-487-826B-13
26	48.2	2.2	6733	4	US-09-124-541-2
27	48	2.2	2787	4	US-09-105-537-40

28	48	2.2	5970	3	US-09-320-878-21	Sequence 21, Appl
29	48	2.2	12730	4	US-09-004-838-91	Sequence 91, Appl
30	47.8	2.2	828	4	US-08-998-416-538	Sequence 538, Appl
31	47.8	2.2	4403765	4	US-09-103-840A-2	Sequence 2, Appl
32	47.6	2.2	1072	1	US-07-971-096-1	Sequence 1, Appl
33	47.6	2.2	1072	1	US-08-175-096-1	Sequence 1, Appl
34	47.6	2.2	168575	1	US-09-426-290-1	Sequence 1, Appl
35	47.4	2.2	3546	1	US-07-951-715A-14	Sequence 14, Appl
36	47.4	2.2	3546	2	US-08-459-448A-14	Sequence 14, Appl
37	47.4	2.2	3546	3	US-08-459-595A-14	Sequence 14, Appl
38	47.4	2.2	3546	3	US-08-459-504B-14	Sequence 14, Appl
39	47.4	2.2	3546	3	US-08-459-444-14	Sequence 14, Appl
40	47.4	2.2	3546	4	US-09-547-422-14	Sequence 14, Appl
41	47	2.1	3468	1	US-07-951-715A-2	Sequence 2, Appl
42	47	2.1	3468	1	US-07-951-715A-4	Sequence 2, Appl
43	47	2.1	3468	2	US-08-459-448A-2	Sequence 4, Appl
44	47	2.1	3468	2	US-08-459-448A-4	Sequence 4, Appl
45	47	2.1	3468	3	US-08-459-595A-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1  
US-08-728-956-3  
; Sequence 3, Application US/08728956  
; Patent No. 5677175  
; GENERAL INFORMATION:  
; APPLICANT: Hodges, Thomas K.  
; APPLICANT: McGee, J. D.  
; TITLE OF INVENTION: Plant Pathogen Induced Proteins  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Barnes & Thornburg  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: USA  
; ZIP: 46204  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/728,956  
; FILING DATE:  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Breen, John P.  
; REGISTRATION NUMBER: 38,833  
; REFERENCE/DOCKET NUMBER: 3220-26119  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (317)231-7745  
; TELEFAX: (317)231-7433  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1591 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Oryza sativa  
US-08-728-956-3

Query Match 4.2%; Score 91.4; DB 1; Length 1591;  
Best Local Similarity 66.9%; Pred. No. 8.4e-12;  
Matches 164; Conservative 0; Mismatches 71; Indels 10; Gaps 2;  
QY 841 ATTACTACCTCCATATTTAATGATGACGCGTTCACITTTTGT---CCACGCTTGAC 897











```

/ TELEFAX: 919-541-8689
/ INFORMATION FOR SEQ ID NO: 186:
/ =====
/ SEQUENCE CHARACTERISTICS:
/     LENGTH: 615 base pairs
/     TYPE: nucleic acid
/     STRANDEDNESS: single
/     TOPOLOGY: linear
/     MOLECULE TYPE: DNA (genomic)
/     ORIGINAL SOURCE:
/         ORGANISM: PAG1074RP
/     US-08-938-416-186

```

Query Match 2.7%; Score 58.4; DB 4; Length 615;  
Best Local Similarity 47.9%; Pred. No. 0.00023;  
Matches 230; Conservative 0; Mismatches 246; Indels

2.

QY	500	TATATATAGGTTTCTFAGCATTACTCAATACACATATATATATATATATATATATCTTA	559
Db	139	TATTATATAAATTTACHTTAATTCACATATATATATATATATATATATATATATATAT -	197
QY	560	GATTCATTATCATATATATATATATAGTAATAGTAAAGCTAGAAAATCTTTAAACCTGAAACAGA	619
Db	198	--TTAAATAGAAATCTATTTAGTCTATGTTCCAAATTTTAAATATAGTTATTTAAATATATAT	255
QY	620	GCGAGTATATATAATTCGGTATATGGAAGTACAGTATTTGTTTATTAATTTGGACGCTTAT	679
Db	256	TAGATATATTATTTTCTTTTAATAAATTTATTAATAGATTTATCAATAAATTAATATATAT	315
QY	680	GAAATACTAGTGAATTCACGCAAACTACATCTTCTTTATTTGGTTATATTTAGTAGTAAT	739
Db	316	TTATTAAATTTGTTTATTAATAATTAATATTTTATATATATAAGAAATTAATTTATTTAAAT	375
QY	740	TTGCTCTCACTAGAAATCTATTTTATCTACAGAGAGATTTAAAATTTATACAAAACCTAAGCTG	799
Db	376	ATTGTAATATATATT - TTTATATTAATATCTATTTTATAAATATTTGTTGATTTATA	434
QY	800	TATACCGGTTGCAACGGAAATAGTAATATGGAAGCCTATCTATTATTTACCTCCATATTTT	859
Db	435	TTATTTAACTTTTATTAAGAAATTAATTATTAATAATTTTAACTTTTAATTTCTTTATAT	494
QY	860	AAATGATGACGCGGTTCACTTTTGTCCACGGTTTGAACCATTCGCTTTATTTAAAAAAAT	919
Db	495	TAAITTTTATATTTATTAATAAATATATATTCATTTTATTTATTTATTTATTTAAAT	554
QY	920	TATGTACGTATATTTATTTATATGACTGTGGTACGTCATCAAAATATTTTTTAAGCATG	979
Db	555	TAAATATTTTAATTAATTTTATCATTTATTAATTAATTAATTAATTAATTAATTAATTAAGATG	614

RESULT 10  
US-09-197-649-7  
; Sequence 7, Application US/09197649  
; Patent No. 6194550  
; GENERAL INFORMATION:  
; APPLICANT: Gold, Larry  
; APPLICANT: Tuerk, Craig  
; APPLICANT: Pribnow, David  
; APPLICANT: Smith, Jonathan D.  
; TITLE OF INVENTION: Systematic Polypeptide Evolution by Reverse Translation  
; FILE REFERENCE: NEX02/C1-CON  
; CURRENT APPLICATION NUMBER: US/09/197,649  
; CURRENT FILING DATE: 1998-11-23  
; EARLIER APPLICATION NUMBER: 07/829,461  
; EARLIER FILING DATE: 1992-01-31  
; EARLIER APPLICATION NUMBER: 07/739,055  
; EARLIER FILING DATE: 1991-08-01  
; EARLIER APPLICATION NUMBER: 07/561,968  
; EARLIER FILING DATE: 1990-08-02  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 390  
; TYPE: DNA

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Sequence
; OTHER INFORMATION: having a 120 repeat of ACG flanked by fixed
; OTHER INFORMATION: fragments having NcoI restriction sites.
US-09-197-649-7

Query Match          2.6%; Score 56.6; DB 4; Length 390;
Best Local Similarity 51.0%; Pred. No. 0.00048;
Matches 134; Conservative 0; Mismatches 129; Indels 0; Gaps

QY 1372 CCACGGCGTATCCGCGCGAGGAGGAGGCCGAGTCCACTACGCACGATPCGACTCGC 1431
      ||| | | | | | | | | | | | | | | | | | | | | |
Db 128 CGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGA 187

QY 1432 CTCCTCATCATCAAGAGGAAGTCACGAGAATCAAGAGAGAGGAGACCGACGTCCGCCA 1491
      ||| | | | | | | | | | | | | | | | | | | |
Db 188 CGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGA 247

QY 1492 COTCAAGGGGACAAGACCOCCTGCTCCGACGTGGCGCCATTGTCGACAAACAAGA 1551
      ||| | | | | | | | | | | | | | | | | |
Db 248 CGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGA 307

QY 1552 CGACGACGCCAGSCCAGSTCAAGGATCTGGAGAGGCCGCTGGCGCGGAGGACGAGGA 1611
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 308 CGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGA 367

QY 1612 CACGGTGCTCACCAAGGTTGTGCA 1634
      || | | | | | | | | | |
Db 368 CCCCATGGTTGGAGTGGTCAGGA 390
```

RESULT 11  
US-08-998-416-288  
Sequence 288, Application US/08998416  
Patent No. 6239264  
GENERAL INFORMATION:  
APPLICANT: Philippsen, Peter  
APPLICANT: Pohlmann, Rainer  
APPLICANT: Steiner, Sabine  
APPLICANT: Mohr, Christine  
APPLICANT: Wendland, Jorgen  
APPLICANT: Knechtle, Philipp  
APPLICANT: Rebischung, Corinne  
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPII  
NUMBER OF SEQUENCES: 1152  
TITLE OF INVENTION: AND USES THEREOF  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NO. 6239264artis Corporation  
STREET: 3054 Cornwallis Road  
CITY: Research Triangle Park  
STATE: NO. 6239264th Carolina  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/998,416  
FILING DATE: 24-DEC-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: CH 0016/97  
FILING DATE: 31-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Weigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8587  
TELEFAX: 919-541-8689





```
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,826B
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CPI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; US-08-487-826B-13

Query Match      2.3%; Score 50.6; DB 2; Length 19124;
Best Local Similarity 44.8%; Pred. No. 0.074;
Matches 236; Conservative 0; Mismatches 289; Indels 2; Gaps 1;

QY 476 TTCCCTGGAAATTCCTCGTTCTATATATATAGAGTTTCTAGCATTTACTCATATTCACAT 535
Db 16231 TTCCACACAGATGTTGGTTTCTTCTAAAAAATAAATGAACATATATGCTGTATGTA 16172

QY 536 ATATATATATATGATATATATCTAGATTCATATATATATATATATATAGGTAATGCTA 595
Db 16171 ATATGATATATATATATATATATCTGTTTCTTAAATGATATATATGATGATGCGTA 16112

QY 596 GAAATCTTAAACCTGAACAGAGGAGATATATATATATGCGTAATGGAACCTAGAGTAT 655
Db 16111 TATATATATATATATATATATATATATATGATGTAATGTAATGTAATGTAATGTA 16052

QY 656 TTGTTTATATATGACCTTTATGAATATCTAGTATAGTATTCACGAACTACATTCCTC 715
Db 16051 ATATATATATATATATATATATATATATGTTGATGTAATGTAATGTAATGTAATG 15992

QY 716 TTTATGTTATATATATATATATATGTTCTACTAGATTCATCTATTTTATCT--AGAGAG 773
Db 15991 TATGATGATATATGTCATGTAATATTTATTTTATTTTATTTTATTTTATTTTATTT 15932

QY 774 ATTTAAATATATTTACAAACCTACCTGATATACCCGTTGCAACGGAATAGTATATGACCT 833
Db 15931 TTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 15872

QY 834 ATTACTTATTTACCTCATATTTTATGATGACGCGTTCTACTTTTCTCCACGTT 893
Db 15871 ATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 15812

QY 894 TGACCATTCCTCTATTTAAATAAATATATATATATATATATATTTTATTTTATGACTGAT 953
Db 15811 AAATTTTATTTTATTTATGATATATTTTATTTTATTTTATTTTATTTTATTTTATTT 15752

QY 954 TCGTCATCAATATTTTATTTTATGACGATGACATTAACATTTTCATATTTG 1000
Db 15751 TTTATGATATATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTG 15705
```

Search completed: March 26, 2003, 13:11:56  
Job time : 246.441 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2003, 05:23:30 ; Search time 121.358 Seconds  
(without alignments)  
15433.627 Million cell updates/sec

Title: US-09-702-134-7212\_COPY\_10200\_12400

Perfect score: 2201

Sequence: 1 cgtcgttgctttgtgcatc.....tagcttgctttgttgta 2201

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications\_NA.\*

1: /cgn2\_6/ptodata/1/pubna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubna/US09\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/1/pubna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubna/US10\_NEW\_PUB.seq.\*  
12: /cgn2\_6/ptodata/1/pubna/US10\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/1/pubna/US60\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/1/pubna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	143.6	6.5	2000	10	US-09-887-576-859
c 2	124.4	5.7	356	10	US-09-770-791-606
c 3	110.2	5.0	689	10	US-09-770-149-308
c 4	100	4.5	33675	10	US-09-921-992-2
c 5	97.4	4.4	2000	10	US-09-887-576-832
c 6	96.6	4.4	2000	10	US-09-887-576-859
c 7	92.8	4.2	2000	10	US-09-887-576-862
c 8	72.6	3.3	33675	10	US-09-921-992-2
c 9	72.4	3.3	2000	10	US-09-887-576-832
c 10	59.6	2.7	550	10	US-09-791-171-53
c 11	56.6	2.6	390	10	US-09-790-399-7
c 12	53.4	2.4	2000	9	US-09-938-842A-4849
c 13	51.6	2.3	2520	9	US-09-858-525A-9
c 14	51.4	2.3	640681	10	US-09-790-988-1
c 15	49.2	2.2	4880	9	US-09-927-827-35
c 16	49	2.2	2000	9	US-09-938-842A-3436
c 17	48.8	2.2	516	10	US-09-960-352-5785
c 18	48.8	2.2	3624	10	US-09-927-112-3
c 19	48.8	2.2	4635	10	US-09-927-112-1

20	48.6	2.2	431	10	US-09-960-352-5558	Sequence 5558, Ap
21	48.6	2.2	1729	9	US-09-938-842A-4225	Sequence 4225, Ap
c 22	48.4	2.2	383	10	US-09-969-373-1100	Sequence 1100, Ap
c 23	48	2.2	2000	9	US-09-938-842A-3641	Sequence 3641, Ap
c 24	48	2.2	2787	9	US-09-860-846-40	Sequence 40, Appl
c 25	48	2.2	2787	10	US-09-861-289-40	Sequence 40, Appl
c 26	47.6	2.2	475	10	US-09-864-761-6203	Sequence 6203, Ap
c 27	47.6	2.2	511	10	US-09-864-761-62817	Sequence 22817, A
c 28	47.6	2.2	2000	9	US-09-938-842A-2717	Sequence 2717, Ap
c 29	47.4	2.2	356	9	US-09-988-462-14	Sequence 14, Appl
c 30	47.4	2.2	302250	10	US-09-962-832-154	Sequence 154, App
c 31	47.4	2.2	513509	9	US-09-754-853A-4	Sequence 4, Appli
c 32	47	2.1	3468	9	US-09-988-462-2	Sequence 2, Appli
c 33	47	2.1	3468	9	US-09-988-462-4	Sequence 4, Appli
c 34	46.6	2.1	267	10	US-09-864-761-27984	Sequence 27984, A
c 35	46.6	2.1	474	10	US-09-864-761-11284	Sequence 11284, A
c 36	45.8	2.1	561	10	US-09-974-300-1212	Sequence 1212, Ap
c 37	45.6	2.1	1835	10	US-09-779-429-2	Sequence 2, Appli
c 38	45.6	2.1	513509	9	US-09-754-853A-4	Sequence 4, Appli
c 39	45.4	2.1	520	9	US-10-184-644-332	Sequence 332, App
c 40	45.4	2.1	659158	9	US-09-771-208-20	Sequence 20, Appl
c 41	45.2	2.1	671	9	US-10-184-644-346	Sequence 346, App
c 42	45	2.0	1118	10	US-09-452-239-37	Sequence 37, Appl
c 43	45	2.0	1146	10	US-09-452-239-3	Sequence 3, Appli
c 44	45	2.0	3809	12	US-10-001-870-68	Sequence 68, Appl
c 45	45	2.0	8537	9	US-09-764-872-724	Sequence 724, App

## ALIGNMENTS

### RESULT 1

US-09-887-576-859/c  
; Sequence 859, Application US/09887576  
; Patent No. US20020144047A1

; GENERAL INFORMATION:

; APPLICANT: Budworth, P.

; APPLICANT: Brown, D.

; APPLICANT: Chang, H.

; APPLICANT: Zhu, T.

; APPLICANT: Han, B.

; APPLICANT: Wang, X.

; APPLICANT: Cooper, Bret

; TITLE OF INVENTION: Promoters for regulation of plant expression

; FILE REFERENCE: 1360.001US1

; CURRENT APPLICATION NUMBER: US/09/887,576

; PRIOR FILING DATE: 2001-06-25

; PRIOR APPLICATION NUMBER: US 60/213,848

; PRIOR FILING DATE: 2000-06-23

; PRIOR APPLICATION NUMBER: US 60/214,087

; PRIOR FILING DATE: 2000-06-23

; PRIOR APPLICATION NUMBER: US 60/258,692

; PRIOR FILING DATE: 2000-12-29

; NUMBER OF SEQ ID NOS: 875

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 859

; LENGTH: 2000

; TYPE: DNA

; ORGANISM: Oryza sativa

US-09-887-576-859

Query Match 6.5%; Score 143.6; DB 10; Length 2000;

Best Local Similarity 79.5%; Pred. No. 5e-24;

Matches 210; Conservative 0; Mismatches 44; Indels 10; Gaps 3;

QY 822 TAATATGGACCTATTACTTACTCTCCATATTTAATGATGACGCGCTTTCACCTTT 881

DB 1592 TAATAGTTTACAAACATATCTCTCTGCTATTTTAAATGATGACGCTTTCACCTTT 1533

QY 882 TTCTCAACGTTTGACCATTCGCTTATTAAAAAAA--TTATGCTATCTATTATTA--TT 938

DB 1532 TTAATAACGTTTGACCTTTCGTTTATTAAAAAATTTTATGATGATTCATCATTTT 1473



QY 939 TTATTGACGTGATTCGTCATCAATATTTTAAAGCAGACATCAATTTTCATATT 998  
 Db 1472 TTATTGACGTGATTTTATCATCAATATTTTAAAGCAGATGATTTTGTATT 1413  
 QY 999 TCAAAAAA-----AAAGCAATGTCATATGTTAGTTCGAAAGTCATGTCATCA 1051  
 Db 1412 TGCATAAAAAATTTGAATAAGACCAATGTCACAAAGCTGGTTAAAAAGTCACGCTGTCA 1353  
 QY 1052 TACATTAATAATACGAGGAGTATT 1075  
 Db 1352 TACATTAATAATGAGGGAGTAT 1329  
 RESULT 2  
 US-09-770-791-606/C  
 ; Sequence 606, Application US/09770791  
 ; Patent No. US200200620141  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Goriach, Jörn  
 ; APPLICANT: An, Yong-Qiang  
 ; APPLICANT: Hamilton, Carol M.  
 ; APPLICANT: Price, Jennifer L.  
 ; APPLICANT: Raines, Tracy M.  
 ; APPLICANT: Yu, Yang  
 ; APPLICANT: Rameaka, Joshua G.  
 ; APPLICANT: Page, Amy  
 ; APPLICANT: Matthew, Abraham V.  
 ; APPLICANT: Ledford, Brooke L.  
 ; APPLICANT: Woessner, Jeffrey P.  
 ; APPLICANT: Haas, William David  
 ; APPLICANT: Garcia, Carlos A.  
 ; APPLICANT: Krieker, Maja  
 ; APPLICANT: Slader, Ted  
 ; APPLICANT: Davis, Keith R.  
 ; APPLICANT: Hoffman, Neil  
 ; APPLICANT: Hoffmann, Patrick  
 ; TITLE OF INVENTION: Expressed Sequences of Arabidopsis  
 ; FILE REFERENCE: 2029 (PARA-018PRV)  
 ; CURRENT APPLICATION NUMBER: US/09/770,791  
 ; PRIOR FILING DATE: 2001-01-26  
 ; PRIOR FILING DATE: 2000-01-27  
 ; NUMBER OF SEQ ID NOS: 999  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 606  
 ; LENGTH: 356  
 ; TYPE: DNA  
 ; ORGANISM: Arabidopsis thaliana  
 ; US-09-770-791-606  
 Query Match 5.7%; Score 124.4; DB 10; Length 356;  
 Best Local Similarity 66.3%; Pred. No. 5.7e-20;  
 Matches 179; Conservative 0; Mismatches 91; Indels 0; Gaps 0;  
 QY 1214 GAGAGGTGAGGAGCAGATCGAGGAGAGCGGAGAGCAGTGGCGAGCAGCAGCCAG 1273  
 Db 284 GAGCAATTAAGCAGATGTTGAAGAGAAAGTAGGCAAACTCTGTTCAGAAACAGAGTCAT 225  
 QY 1274 CTGTGAGAGAGTGGAGTGGCGCTCTCCGCCCGGTGGCGAGCTCAGCCGAGGTCCC 1333  
 Db 224 CTGTGAGAGAGTGTGATGAAGACTCTCTGTCTGTGTGGAGAGTTGGTAAGGCCCT 165  
 QY 1334 AGAATTCAGGTGGAGAGACTCTCTTCACCAAGCCGAGCGGTCATCCGCGCGAG 1393  
 Db 164 AGAATTCAGAGTGTGAGTGTGATGTTTACAAAGAGCAGTGTGTGTGTGTGTGTGT 105  
 QY 1394 GAGGAGCGGAGTCCACCTACGAGCAGCTCGCTCTCCATCATCAAGAGGAG 1453  
 Db 104 GAAGATGCTGAGACAGTATAGCTTGTATCGACTTGGTATCAACGATAATACAGAGGAG 45  
 QY 1454 CTCAGGAAGATCAAGGAGAGGAGCCGAC 1483

Db 44 CTGAGGAAGATCAAGGAGAGGAGTCAAGAC 15  
 RESULT 3  
 US-09-770-149-308  
 ; Sequence 308, Application US/09770149  
 ; Patent No. US20020059663A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Goriach, Jörn  
 ; APPLICANT: An, Yong-Qiang  
 ; APPLICANT: Hamilton, Carol M.  
 ; APPLICANT: Price, Jennifer L.  
 ; APPLICANT: Raines, Tracy M.  
 ; APPLICANT: Yu, Yang  
 ; APPLICANT: Rameaka, Joshua G.  
 ; APPLICANT: Page, Amy  
 ; APPLICANT: Matthew, Abraham V.  
 ; APPLICANT: Ledford, Brooke L.  
 ; APPLICANT: Woessner, Jeffrey P.  
 ; APPLICANT: Haas, William David  
 ; APPLICANT: Garcia, Carlos A.  
 ; APPLICANT: Krieker, Maja  
 ; APPLICANT: Slader, Ted  
 ; APPLICANT: Davis, Keith R.  
 ; APPLICANT: Hoffman, Neil  
 ; APPLICANT: Hoffmann, Patrick  
 ; TITLE OF INVENTION: Expressed Sequences of Arabidopsis  
 ; FILE REFERENCE: 2024 (PARA-013PRV)  
 ; CURRENT APPLICATION NUMBER: US/09/770,149  
 ; PRIOR FILING DATE: 2001-01-26  
 ; PRIOR FILING DATE: 2000-01-27  
 ; NUMBER OF SEQ ID NOS: 999  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 308  
 ; LENGTH: 689  
 ; TYPE: DNA  
 ; ORGANISM: Arabidopsis thaliana  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)...(689)  
 ; OTHER INFORMATION: n = A,T,C or G  
 ; US-09-770-149-308  
 Query Match 5.0%; Score 110.2; DB 10; Length 689;  
 Best Local Similarity 62.2%; Pred. No. 2.2e-16;  
 Matches 163; Conservative 0; Mismatches 99; Indels 0; Gaps 0;  
 QY 1214 GAGAGGTGAGGAGCAGATCGAGGAGAGCGGAGAGCAGTGGCGAGCAGCAGCCAG 1273  
 Db 362 GAGCAATTAAGCAGATGTTGAAGAGAAAGTAGGCAAACTCTGTTCAGAAACAGAGTCAT 421  
 QY 1274 CTGTGAGAGAGTGGAGTGGCGCTCTCCGCCCGGTGGCGAGCTCAGCCGAGGTCCC 1333  
 Db 422 CTGTGAGAGAGTGTGATGAAGACTCTCTGTCTGTGTGGAGAGTTGGTAAGGCCCT 481  
 QY 1334 AGAATTCAGGTGGAGAGACTCTCTTCACCAAGCCGAGCGCTCATCCGCGCGAG 1393  
 Db 482 AGAATTCAGAGTGTGAGTGTGATGTTTACAAAGAGCAGTGTGTGTGTGTGTGTGT 541  
 QY 1394 GAGGAGCGGAGTCCACCTACGAGCAGCTCGCTCTCCATCATCAAGAGGAG 1453  
 Db 542 GAAGATGCTGAGACAGTATAGCTTGTATCGACTTGGTATCAACGATAATACAGAGGAG 601  
 QY 1454 CTCAGGAAGATCAAGGAGAGGAG 1475  
 Db 602 CTGAGGAAGANNNNNNNNNN 623  
 RESULT 4

US-09-921-992-2/c

```
; Sequence 2, Application US/09921992
; Patent No. US20020069426A1
; GENERAL INFORMATION:
; APPLICANT: Boronat, Albert;
; APPLICANT: Campos, Narciso;
; APPLICANT: Rodriguez-Concepcion, Manuel;
; APPLICANT: Seeman, Myriam;
; APPLICANT: Valentín, Henry E.;
; APPLICANT: Venkatesh, Tyamagondlu V.;
; APPLICANT: Venkatesh, Mylavaraou
; TITLE OF INVENTION: Methyl-D-Erythritol Phosphate Pathway Genes
; FILE REFERENCE: 16516.107/35-21(51897)US
; CURRENT APPLICATION NUMBER: US/09/921,992
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/223,483
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 85
; SEQ ID NO 2
; LENGTH: 33675
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6924)..(7019), (7163)..(7269), (7344)..(7444), (7525)..(7634),
; LOCATION: (7694)..(7813), (7923)..(8153), (8253)..(8369), (8515)..(8583),
; LOCATION: (9012)..(9071), (9163)..(9225), (9328)..(9472), (9589)..(9730),
; LOCATION: (9951)..(10028), (10134)..(10293), (10694)..(10798),
; LOCATION: (11028)..(11129)
; NAME/KEY: unsure
; LOCATION: (1..33675)
; OTHER INFORMATION: unsure at all n locations
US-09-921-992-2
```

```
Query Match 4.5%; Score 100; DB 10; Length 33675;
Best Local Similarity 67.6%; Pred. No. 9.5e-13;
Matches 173; Conservative 0; Mismatches 75; Indels 8; Gaps 2;

QY 843 TACTACCTCCATATTTTATGATGACGCGCTTCACCTTTTCTCCACGTTTGACCATTC 902
DB 16817 TACTCCCTCCGTTCTTAAATATGATGCGGTGACCTTTTAAACAGTTTGACCGTTC 16758

QY 903 GTCTATTAAAAAATATGATCTATCTATTTATTTATTTATGATGATGATCTATCA 962
DB 16757 GTCTATTAAAAAATTTAAATAAGTATATTT-TTTTCTATCATTTGATGATGTTA 16699

QY 963 AATATTTTAAAGCATGACATTAACATTTTCATATTTGCAAAA-----AAAAAAG 1015
DB 16698 AATATCTTTTATGATACATATAGTTTATATATTTTCAAAAAATTTTGAATAAGATG 16639

QY 1016 AATGTCAAATGTTAGTCGAAAAGTCAATGCTGTCATACATTAATAATACGAGGATAT 1075
DB 16638 AATGTCAAACAGTCTTAAAGTAAAGTAAAGTAAATTTTGAACGAGGAGTAT 16579

QY 1076 ATGGGATGATCAGT 1091
DB 16578 ATCCGTATCTATCGT 16563

RESULT 5
US-09-887-576-832/c
; Sequence 832, Application US/09887576
; Patent No. US20020144047A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: Promoters for regulation of plant expression
```

```
; FILE REFERENCE: 1360.001USI
; CURRENT APPLICATION NUMBER: US/09/887,576
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 832
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-887-576-832

Query Match 4.4%; Score 97.4; DB 10; Length 2000;
Best Local Similarity 63.1%; Pred. No. 5.4e-13;
Matches 185; Conservative 0; Mismatches 101; Indels 7; Gaps 2;

QY 820 AGTAATATGAGACCTATTACTACCTCCATTTTAATGATGATGACGCGTCTACT 879
DB 1781 ATTAATTTTGGACCCGAGAGAGTACTTCTCCGTCCTAAATATTGACACCATGATT 1722

QY 880 TTTTGTCCACGCTTTGGACCATCTCTTATTAAATAAAATATCTATCTATTTATTT 939
DB 1721 TTTTAAACATGTTTGACCGTCTCTTATCAAAAATTAAGTATTAATTAATTT-TTT 1653

QY 940 TATTATGACTTGATCTGTCATCAATATTTTAAAGCATGACATCAACATTTTCATATT 999
DB 1662 TTCTATCATTTGATTCATTTAAATATATTTTCCGTATATATATAGTTTCATATTTTC 1603

QY 1000 GCAAAA-----AAAAAAGCAATGCTCAATGTTAGTCGAAAAGTCAATGCTGCATA 1053
DB 1602 ACAAAAGTTTAAATAAGCAATGCTCAAAATGTTTAAAGTCAACACGCTCA 1543

QY 1054 CATTAAATATGAGGAGGATATATGAGTATGATGATGATGATGATGATGATGATGAT 1106
DB 1542 TATTAGGACGAGGAGGATATATGATGATGATGATGATGATGATGATGATGATGAT 1490

RESULT 6
US-09-887-576-859
; Sequence 859, Application US/09887576
; Patent No. US20020144047A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360.001USI
; CURRENT APPLICATION NUMBER: US/09/887,576
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 859
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-887-576-859

Query Match 4.4%; Score 96.6; DB 10; Length 2000;
```

```

Best Local Similarity 67.1%; Pred. No. 8.4e-13;
Matches 171; Conservative 0; Mismatches 74; Indels 10; Gaps 2;

QY 834 ATTACTATTATTACCTCCCATATTTTAAATGATGACGCGGTTCACTTTTGTGTCACACGTT 893
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1321 ATTATGTAATACCTCCCTCCCATATTTTAAATGATGACACGCGTTGACTTTTAAACACGTT 1380

QY 894 TGACCATTCGCTGTTATTAATAAAAAAATATGATGACT-----ATTATTATTATTATTATGAC 948
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1381 TGACCATTCGCTGTTATTAATAAAAAAATATGATGACT-----ATTATTATTATTATTATGAC 1440

QY 949 TTGATTGCTGATCAAAATATTTTAAAGCATGACATATAACATTTTCATATTTTGCAA----- 1003
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1441 AGAACATTGATGATTAATCAAGTCACATTAATAAATAATGATGATTAATACATAAATTTT 1500

QY 1004 AAAAAAAAACGAATGGTCAATGTTAGTGGAAAGTCAATGGTGTCAATGATTAATAATA 1063
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1501 TTATAATAACGAAGAGTCGAAAGCTTTTATTAATAAAGTCACTGCGTTATACATTAATAATA 1560

QY 1064 CGGAGGGATTATATG 1078
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1561 CGAGGAAGTATATG 1575

RESULT 7
US-09-887-576-862
; Sequence 862, Application US/09887576
; Patent No. US20020144047A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360.001US1
; CURRENT APPLICATION NUMBER: US/09/887,576
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/259,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 862
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-887-576-862

Query Match 4.2%; Score 92.8; DB 10; Length 2000;
Best Local Similarity 67.4%; Pred. No. 6.8e-12;
Matches 163; Conservative 0; Mismatches 72; Indels 7; Gaps 2;

QY 843 TACTACTCCCATATTTTAAATGATGACGCGGTTCACTTTTGTGTCACACGTTGACCAT 901
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1087 TACTCCTCTGTTTCAATAACTGTTATATTTTAAATATGTTTGTGATCGTT 1146

QY 902 CGTCTTATTAATAAAATATCTATCTATTTATTTTATTTATGATGATGATGATGATGATC 961
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1147 CGTCTTATCTAATAATTAAGTAATTAATTCCTTTCCCTATCATTTGATTCATCGTT 1206

QY 962 AAATATTTTAAAGCATGACATAAATATTTTCATTTTCAATTTGCAAAA-----AAAAAAG 1015
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1207 AAATATGATGATGATACATATAGTTTTCATATTTTACATAATTTTCAAAAAGTTTGTATGATGAC 1266

QY 1016 AATGTCGAATGTTAGTCGAAAGTCAATGGTGTCAATCAATTAATAATACGAGGGAATAT 1075
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1267 AACGCTCAATATGCTATAAAGTCAACGGTATCAAAATATTTTAAAGGAGGGGATAT 1326
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360.001051
; CURRENT APPLICATION NUMBER: US/09/887,576
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 832
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-887-576-832

Query Match      3.3%; Score 72.4; DB 10; Length 2000;
Best Local Similarity 58.3%; Pred. No. 5.1e-07;
Matches 148; Conservative 0; Mismatches 101; Indels 5; Gaps 1;

QY 826 ATGACGATTTACTTATCTACCTCCATATTTTAAATGATGACCGCTTCACTTTTGT 885
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1505 ACGGATGCTACTACATATACCTCTCCGCTCCCTAAATATATGACGTTGTTAA 1564

QY 886 CCACGTTTGACCATCTGCTTATTAATAAATAATGATGATATATTTATTTAT 945
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1565 AACATTTTGACCATCTGCTTATTAATAAATAATGATGATATATTTATTTAT 1624

QY 946 GACTTGATTCGTCATCAATATTTTAAAGCATGACATAAATTTTCAATTTTCAA-- 1003
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1625 ACGGAAATATATTTATATGATCAATGATAGAAAAATATTTATTTACTTAAT 1684

QY 1004 ---AAAAAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1060
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1685 TTTTGAATAGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1744

QY 1061 ATACGAGGAGATTA 1074
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1745 GCACGAGGAGATTA 1758

RESULT 10
US-09-791-171-63
; Sequence 63, Application US/09/91171
; Patent No. US20020094336A1
; GENERAL INFORMATION:
; APPLICANT: ANDERSEN, Peter
; APPLICANT: NIELSEN, Rikke
; APPLICANT: OETTINGER, Thomas
; APPLICANT: RASMUSSEN, Peter Birk
; APPLICANT: ROSENKRANDS, Ida
; APPLICANT: WELDLINGH, Karin
; APPLICANT: FLORIO, Walter
; TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
; FILE REFERENCE: 670001-2002.1
; CURRENT APPLICATION NUMBER: US/09/791,171
; CURRENT FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 09/050,739
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 0376/97
; PRIOR FILING DATE: 1997-04-02
; PRIOR APPLICATION NUMBER: 1277/97
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/044,624

```

```

; PRIOR FILING DATE: 1997-04-18
; PRIOR APPLICATION NUMBER: 60/070,488
; PRIOR FILING DATE: 1998-01-05
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 63
; LENGTH: 550
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-09-791-171-63

Query Match      2.7%; Score 59.6; DB 10; Length 550;
Best Local Similarity 50.3%; Pred. No. 0.00024;
Matches 146; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

QY 1375 CGGCGTCATCCGCGCGGAGGAGACGCGAGTCCACCTACCCAGCATGACCTCGCCTC 1434
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 250 CGGTGCGCGCGTCCGAGGCTCCGAGGAGCAGTCCGAGTTCGACGTGATCCTTCAGGCGGC 309

QY 1435 CTCATCATCAAGAGGAAGCTCAGGAAGATCAAGGAGAGGAGAGCCGCTCGGCCACT 1494
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 310 CGCGGACAAAGATCGCGCTCATCAAGGTGTCGCGGAGATCGTTTCGCGCTGGGCT 369

QY 1495 CAAGGAGGACCAAGCCCGCTCCGACTGGCGCCATCATCTGCTCGACAAACAGAGA 1554
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 370 CAAGGAGGACCAAGGACCTGTCGACGCGCGCCCAAGCCGCTGCTGGAGAGGTGCCAA 429

QY 1555 CGACGCCAGGCGCCAGCTCAAGGATCTGGAGGAGCGCGCTCGCGCGGAGGAGGAGAC 1614
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 430 GGAGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 489

QY 1615 GGTGCTCACCAGGCTGTCAGTCTGATGATGATGATGATGATGATGATGATGATGAT 1664
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 490 GTAGCTCTCCAGCGGTGTTCTTTTGGTCTGCTGCGCGCGGAGGAGGAGGAGGAGGAG 539

RESULT 11
US-09-790-399-7
; Sequence 7, Application US/09/90399
; Patent No. US20020038000A1
; GENERAL INFORMATION:
; APPLICANT: Gold, Larry
; APPLICANT: Tuerk, Craig
; APPLICANT: Priebow, David
; APPLICANT: Smith, Jonathan D.
; TITLE OF INVENTION: Systematic Polypeptide Evolution by Reverse Translation
; FILE REFERENCE: NEX02/CI-CON2
; CURRENT APPLICATION NUMBER: US/09/790,399
; CURRENT FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: 09/197,649
; PRIOR FILING DATE: 1998-11-23
; PRIOR APPLICATION NUMBER: 07/829,461
; PRIOR FILING DATE: 1992-01-31
; PRIOR APPLICATION NUMBER: 07/739,055
; PRIOR FILING DATE: 1991-08-01
; PRIOR APPLICATION NUMBER: 07/561,968
; PRIOR FILING DATE: 1990-08-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 390
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Sequence
; OTHER INFORMATION: having a 120 repeat of ACG flanked by fixed
; OTHER INFORMATION: fragments having NcoI restriction sites.
US-09-790-399-7

Query Match      2.6%; Score 56.6; DB 10; Length 390;
Best Local Similarity 51.0%; Pred. No. 0.00096;
Matches 134; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

```

QY 1372 CCACGGCGTCTCCGCGGAGGAGGAGCGCGGAGTCCACCTACCGCCAGCATCGACCTGCG 1431  
 Db 128 CGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 187  
 QY 1432 CTCCTCCATCATCAAGAGGAGCTCAGGAAGATCAAGGAGAGGAGACGCGTCCGCCA 1491  
 Db 188 CGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 247  
 QY 1492 CCTCAAGGGGACCAAGCCCGCTCTCCGAGCTGGCGGCCCATCTTCTCGCAACAACGA 1551  
 Db 248 CGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 307  
 QY 1552 CGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1611  
 Db 308 CGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 367  
 QY 1612 CACGCTGCTCACCAAGTCTGCA 1634  
 Db 368 CGCCATGTTGCGATGTCAGGA 390  
 RESULT 12  
 US-09-938-842A-4849/C  
 ; Sequence 4849, Application US/09938842A  
 ; Patent No. US20020160378A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Harper, Jeff  
 ; APPLICANT: Kreps, Joel  
 ; APPLICANT: Wang, Xun  
 ; APPLICANT: Zhu, Tong  
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
 ; FILE REFERENCE: SRIPI300-3  
 ; CURRENT APPLICATION NUMBER: US/09/938.842A  
 ; CURRENT FILING DATE: 2001-08-24  
 ; PRIOR APPLICATION NUMBER: US 60/727,866  
 ; PRIOR FILING DATE: 2000-08-24  
 ; PRIOR APPLICATION NUMBER: US 60/264,647  
 ; PRIOR FILING DATE: 2001-01-16  
 ; PRIOR APPLICATION NUMBER: US 60/300,111  
 ; PRIOR FILING DATE: 2001-06-22  
 ; NUMBER OF SEQ ID NOS: 5379  
 ; SEQ ID NO 4849  
 ; LENGTH: 2000  
 ; TYPE: DNA  
 ; ORGANISM: Arabidopsis thaliana  
 US-09-938-842A-4849  
 Query Match 2.4%; Score 53.4; DB 9; Length 2000;  
 Best Local Similarity 48.3%; Pred. No. 0.018;  
 Matches 184; Conservative 0; Mismatches 191; Indels 6; Gaps 1;  
 QY 626 ATATATPAATTCGGTAATGGAACTAGAGTATTGTTTATTAATTTGACCTTTATGAATA 685  
 Db 1923 ATGTGATTTACAGGGAATTAATGTTTATGTCACAAACAAAATAATCAATGATTTA 1864  
 QY 686 CTAGTAGAATTCACGAAATACATCTCTTTATGTTATGTTATATGTTAGTACTGTTGTT 745  
 Db 1863 TTGAAAAATGCAAACTTCAGAGTTTATTATCCAAATATTTTATGTTGAATAATATTTT 1804  
 QY 746 CACTAGATTTCTATTATATCTAGAGATTAATAATTTATACAAATACACTGTATACC 805  
 Db 1803 CTGTTCTTTATTTATTTATTTTATGAAAAAATTTGTTTCAAAAGATTTATTTTAAAA 1744  
 QY 806 CGTGCACGGAAGAATATGACCTATTACTTATTTACTTACCTCCATATTTTATGTA 865  
 Db 1743 AGTTTATGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1690  
 QY 866 TGACGCGGTCACCTTTGTTGCCAAGTTTGACCATTCGTTCTATTAATAAAATATGTA 925  
 Db 1689 AATAATGAATATTTAGTTTAAAGTTATATAAAAAATATAAATACAAAAATTAATATA 1630  
 QY 926 TCATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 985

Db 1629 TTATGATTAGTTTATTAATATATTTTATTTATTTATTTATTTATTTATTTATTTATTT 1570  
 QY 986 ACAATTTTCATATTTGCAAAA 1006  
 Db 1569 AATTTTAATAGATTTAAATAA 1549  
 RESULT 13  
 US-09-858-525A-9  
 ; Sequence 9, Application US/09858525A  
 ; Publication No. US20020199215A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BOETS, Anemie  
 ; APPLICANT: ARNAUT, Greta  
 ; APPLICANT: VAN RIE, Jeroen  
 ; APPLICANT: DAMME, Nicole  
 ; TITLE OF INVENTION: No. US20020199215alel Toxins  
 ; FILE REFERENCE: 021565-077  
 ; CURRENT APPLICATION NUMBER: US/09/858.525A  
 ; CURRENT FILING DATE: 2001-05-17  
 ; PRIOR APPLICATION NUMBER: US 09/573,872  
 ; PRIOR FILING DATE: 2000-05-18  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 9  
 ; LENGTH: 2520  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: modified DNA encoding ISPLA protein  
 ; NAME/KEY: CDS  
 ; LOCATION: (3)..(2510)  
 US-09-858-525A-9  
 Query Match 2.3%; Score 51.6; DB 9; Length 2520;  
 Best Local Similarity 46.4%; Pred. No. 0.056;  
 Matches 168; Conservative 0; Mismatches 194; Indels 0; Gaps 0;  
 QY 1266 ACAGCCAGCTTGTGAGGAGTGGAGTGGCGCTCTCGCGCGGCTGGCGAGCTCAGCC 1325  
 Db 157 ACACCTCTGTGACCAAGACACAGGAGTACCCTCCATCGCTGGATCGGCTGATCC 216  
 QY 1326 GAGGTCCTCAAGATTTGAGTGGAGATCATCTCTTCAACAGCGCCAGCGGTGATCC 1385  
 Db 217 AGTCTCGCGCACTGGTACTTCACCTTCAAGCTCGGACGAGAGACGCAATCATCG 276  
 QY 1386 GCGCGGAGGAGCGCGAGTCCACCTACGCGAGCATGACCTCGCTCTCTCATCATCA 1445  
 Db 277 AGCTGGACGCAAGTATCTCGAGAGGGACACACAGAGTCCGTGACCTCGAAA 336  
 QY 1446 AGAGGAAGCTCAGGAAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1505  
 Db 337 AGGGCCAGCTGGTGCAGATCAAGATCGAGTACCAGTCCGACGAGCGCTGACATCGACA 396  
 QY 1506 AGCCCCCGTCTCCGATGCGCCCATCATTTGTCGACACACAGAGAGAGAGAGAGAG 1565  
 Db 397 ACAAGATCTTCAAGAGAGCTGAAGCTGTTCAAGATGAGTCTCCAGAGACCACTCC 456  
 QY 1566 CCAGCTCAAGGATCTGAGAGGCGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1625  
 Db 457 TGACGAGGACGAGCTGAGGAGCCAGAGTTCACACAGAGGAGAGCCAGGTTCTCTGA 516  
 QY 1626 AG 1627  
 Db 517 AG 518  
 RESULT 14  
 US-09-790-988-1  
 ; Sequence 1, Application US/09790988  
 ; Patent No. US20020127687A1  
 ; GENERAL INFORMATION:



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2003, 18:27:39 ; Search time 361.22 Seconds  
(without alignments)

14774.266 Million cell updates/sec

Title: US-09-702-134-7212\_COPY\_28000\_30500

Perfect score: 2501

Sequence: 1 atgtccagataaaattggc.....ctatcaatggttgcctcaat 2501

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125995159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Lasting first 45 summaries

Database : N\_Geneseq\_101002.\*

1: /SID82/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SID82/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SID82/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SID82/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SID82/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SID82/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SID82/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SID82/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SID82/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SID82/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SID82/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SID82/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SID82/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SID82/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SID82/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SID82/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SID82/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SID82/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SID82/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SID82/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SID82/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SID82/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SID82/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SID82/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	292.4	11.7	1259	16	AAQ75360
2	292.4	11.7	1259	22	AAQ75360
3	290.8	11.6	1169	21	AAQ75361
4	271	10.8	1490	16	AAQ75361
5	229.6	9.2	1420	24	AAQ75361
6	228	9.1	1395	16	AAQ75359
7	226.4	9.1	1161	21	AAQ75359
8	174	7.0	1122	22	AAQ75359
9	153.2	6.1	1077	21	AAQ75359

10	129	5.2	811	21	AAZ99478
11	126.4	5.1	2112	22	AAH46558
12	114.8	4.6	1187	22	AAH46555
13	94.4	3.8	799	19	AAV55831
14	94.4	3.8	1926	21	AAH50254
15	94.4	3.8	1926	22	AAH50254
16	94.4	3.8	2580	21	AAH75454
17	94.4	3.8	2580	24	AAH75454
18	94.4	3.8	5452	20	AAH90923
19	94.4	3.8	8705	20	AAZ23778
20	94.4	3.8	9600	19	AAV21683
21	94.4	3.8	10380	20	AAZ22248
22	94.4	3.8	10596	14	AAQ51731
23	94.4	3.8	10596	17	AAH40348
24	94.4	3.8	10596	20	AAH15650
25	94.4	3.8	16080	21	AAH59553
26	94.2	3.8	1181	20	AAZ27502
27	94.2	3.8	1181	22	AAH09902
28	94	3.8	114793	22	AAH08215
29	92.8	3.7	1925	20	AAH90924
30	92	3.7	8298	22	AAH72613
31	89	3.6	7720	21	AAH53800
32	85.8	3.4	1000	21	AAH02484
33	84.4	3.4	795	19	AAH55830
34	83.6	3.3	1948	22	AAH46557
35	79.8	3.2	1505	15	AAQ55750
36	79.2	3.2	2744	16	AAQ98470
37	77.2	3.1	672	24	ABL59991
38	76.2	3.0	1274	21	AAH33593
39	75.4	3.0	174	16	AAQ75364
40	74.2	3.0	1105	20	AAH04613
41	74.2	3.0	1690	20	AAH04614
42	74.2	3.0	114955	20	AAH53491
43	73.2	2.9	276	24	ABL75810
44	73	2.9	446	22	AAH50433
45	73	2.9	446	22	AAH68382

# ALIGNMENTS

RESULT 1	AAQ75360	standard; cdna to mRNA; 1259 BP.
XX	AAQ75360;	
AC	AAQ75360;	
XX		
DT	01-JUL-1995 (first entry)	
DE	Gibberellin-20-oxidase.	
XX	Gibberellin-20-oxidase; plasmid pat2301; transgenic plant;	
KW	crop improvement; ss.	
XX		
OS	Arabidopsis thaliana.	
XX		
FE	Key	Location/Qualifiers
FT	CDS	12..1142
FT		/*tag= a
FT		/note= "GA-20-oxidase"
XX		
PN	W09428141-A.	
XX		
PD	08-DEC-1994.	
XX		
PF	24-MAY-1994; 94WO-EP01664.	
XX		
PR	28-MAY-1993; 93GB-0011147.	
XX		
PA	(LONG-) LONG ASHTON RES STATION.	
XX		
PI	Graebe JE, Hedden P, Lange T, Phillips A;	
XX		







PR	21-JUN-1999;	9905-0139817;
PR	22-JUN-1999;	9905-0139899;
PR	23-JUN-1999;	9905-0140353;
PR	24-JUN-1999;	9905-0140354;
PR	25-JUN-1999;	9905-0140695;
PR	28-JUN-1999;	9905-0140823;
PR	30-JUN-1999;	9905-0140991;
PR	30-JUN-1999;	9905-0141287;
PR	01-JUL-1999;	9905-0141842;
PR	01-JUL-1999;	9905-0142154;
PR	02-JUL-1999;	9905-0142055;
PR	06-JUL-1999;	9905-0142390;
PR	08-JUL-1999;	9905-0142803;
PR	09-JUL-1999;	9905-0142920;
PR	12-JUL-1999;	9905-0142977;
PR	13-JUL-1999;	9905-0143542;
PR	14-JUL-1999;	9905-0143624;
PR	15-JUL-1999;	9905-0144005;
PR	16-JUL-1999;	9905-0144086;
PR	16-JUL-1999;	9905-0144086;
PR	19-JUL-1999;	9905-0144325;
PR	19-JUL-1999;	9905-0144331;
PR	19-JUL-1999;	9905-0144332;
PR	19-JUL-1999;	9905-0144333;
PR	19-JUL-1999;	9905-0144334;
PR	19-JUL-1999;	9905-0144335;
PR	20-JUL-1999;	9905-0144352;
PR	20-JUL-1999;	9905-0144632;
PR	20-JUL-1999;	9905-0144684;
PR	21-JUL-1999;	9905-0144814;
PR	21-JUL-1999;	9905-0145086;
PR	22-JUL-1999;	9905-0145088;
PR	22-JUL-1999;	9905-0145085;
PR	22-JUL-1999;	9905-0145087;
PR	22-JUL-1999;	9905-0145089;
PR	22-JUL-1999;	9905-0145192;
PR	23-JUL-1999;	9905-0145143;
PR	23-JUL-1999;	9905-0145218;
PR	23-JUL-1999;	9905-0145224;
PR	26-JUL-1999;	9905-0145276;
PR	27-JUL-1999;	9905-0145913;
PR	27-JUL-1999;	9905-0145918;
PR	27-JUL-1999;	9905-0145919;
PR	28-JUL-1999;	9905-0145951;
PR	02-AUG-1999;	9905-0146386;
PR	02-AUG-1999;	9905-0146388;
PR	02-AUG-1999;	9905-0146389;
PR	03-AUG-1999;	9905-0147036;
PR	04-AUG-1999;	9905-0147204;
PR	04-AUG-1999;	9905-0147302;
PR	05-AUG-1999;	9905-0147192;
PR	05-AUG-1999;	9905-0147260;
PR	06-AUG-1999;	9905-0147303;
PR	09-AUG-1999;	9905-0147416;
PR	09-AUG-1999;	9905-0147493;
PR	09-AUG-1999;	9905-0147935;
PR	10-AUG-1999;	9905-0148171;
PR	11-AUG-1999;	9905-0148319;
PR	12-AUG-1999;	9905-0148341;
PR	13-AUG-1999;	9905-0148565;
PR	13-AUG-1999;	9905-0148668;
PR	16-AUG-1999;	9905-0149375;
PR	17-AUG-1999;	9905-0149175;
PR	18-AUG-1999;	9905-0149426;
PR	20-AUG-1999;	9905-0149722;
PR	20-AUG-1999;	9905-0149722;
PR	20-AUG-1999;	9905-0149929;
PR	23-AUG-1999;	9905-0149902;
PR	23-AUG-1999;	9905-0149930;
PR	25-AUG-1999;	9905-0150566;
PR	26-AUG-1999;	9905-0150684;
PR	27-AUG-1999;	9905-0151065;
PR	27-AUG-1999;	9905-0151066;

PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	13-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	23-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Query Match 11.6%; Score 290.8; DB 21; Length 1169;

Query Match 11.6%; Score 290.8; DB 2.4e-48;  
Best Local Similarity 59.9%; Pred. No 2.4e-48;

Best Local Similarity 39.9%; Pseud. NO: 2.4e-48  
Matches 531; Conservative 0; Mismatches 337

2;

Qy	939	GAGGCGTGGAGAGCAGCGCTTCTTCGTGTGTGTTAAACACAGC	998
Db	288	GAGGCGCTTAAGAAGCAGCGTTCTTCGTGTGTGTTCAATACAGC	347
Qy	999	CTGGAGGAGGCGCACCGGTGCATGAGCGCTTCTTCACGCTGCG	1058
Db	348	ATTTCAGCGCTCATGAATACAGAGCGCTCTTTGATATGCTCTCT	407
Qy	1059	CGGCGCAGCGCGCGCGGGGAGAGCTCGCGCTACGCGACAGCTTC	1118
Db	408	AGGGTCTTTAGAAATCCGCTGAGAGTGTGGCTACGCAAGCAGTT	467
Qy	1119	GCCTCCACAGCTGCGGTGGAAGAGACGCTGCTTCGGTACTCATG	1178
Db	468	TCACACAGCTTCCATGGAAGAGACCCCTTCTTCGGT-----	518
Qy	1179	GAGGCGCAGGAGGCGGTGGGTGAGTACTGTTGGAGAGTCGCGG	1238
Db	519	ATGAGCGCGCTCAAAATCCGTTCAAGATTACTTCTCGCATGGTTG	578



Db	814	GAOCTCATGTGATCCAAAGTCTTTTGGACCATCTTTCATCAAGACCATGTCAATGGCCTTC	873
QY	1498	AGGTGTGGCGGAGGGCGGTGGCGCGCATCGCCCTCGCCCGCGGGCGCTGCTGCTCA	1557
Db	874	AAAGTCTTTGTGCACATCAATGCGCATCCATTCGTCCGATCCCAAGGCTTTGCTGTGCA	933
QY	1558	ACGTGGCGGACACCTTCATGGGCGCTCCCAAGCGCAGGTACCGCAGCTGCCTGCACCGGG	1617
Db	934	ATAATGGTGCACACTTCATGGCTCTATCGAAGCGGATATTCAGAGCTGTTTGCATAGAG	993
QY	1618	CGGTGCTCAACAGCAGCGGCCCTCGCGCGCTCGCTGGCTTCTTCTCTGCCCGGAGATGG	1677
Db	994	CGGTGTGAATAGAGAGAGCGCGAGAAATCGATGCGCTTTTCTGTGTCCGAAAGAAG	1053
QY	1678	ACAGGTGTGTGGCGCGCGGAGGAGCTGTGCGAGCACCACCGCAGGAGGTGTACCCGG	1737
Db	1054	ACAAGTGTGTGAACCAACCAAGTGATATTTTGGAGAGTGAACAAGAAAAATACCCCTG	1113
QY	1738	ACTTCAAGTGGCGGCGCTCTCTGGACTTCAGCGAGCGCCACTACAGGCGCGACATGCGCA	1797
Db	1114	ACTTCACTTGGCTATGTCTCTTGAGTTCACATAAACAATACCGAGCAGATGTGAATA	1173
QY	1798	CGCTTACGCTTCTTCGACTGCTTAAT	1826
Db	1174	CTCTCGATTCCCTTTCTAATGGGTATT	1202
RESULT 5			
AAD21991			
XX	ID	AAD21991 standard; cDNA; 1420 BP.	
XX	AC		
XX	AC		
DT	12-FEB-2002	(first entry)	
XX	Citrullus lanatus	Cv20ox cDNA.	
DE	XX		
KW	Testa specific Cv20ox gene; gibberellin 20-oxidase; GA 20-oxidase;		
KW	Cv20ox promoter; GA20oxP; seedless fruit generation; watermelon;		
KW	melon; SS.		
OS	Citrullus lanatus.		
XX	Key	Location/Qualifiers	
FX	FT	70..1209	
FT	ET	/*tag= a	
FT	FT	/product= "C. lanatus GA 20-oxidase"	
XX	WO200177310-A1.		
PN	XX		
PD	18-OCT-2001.		
XX	09-OCT-2000; 2000WO-KR01127.		
XX	08-APR-2000; 2000KR-0018483.		
XX	(DONG-) DONGBU HANNONG CHEM CO LTD.		
XX	An G, Kang H, Jun S, Kim J, Jung K, Park Y, Lee S;		
PI	WPI; 2002-010905/01.		
XX	DR	P-PSDB; AAE13273.	
XX	Novel testa specific Cv20ox gene, a gibberellin 20-oxidase gene, which		
PT	is integument-specifically expressed in seeds of Citrullus lanatus		
PT	useful for producing seedless fruits such as melon		
XX	Claim 2; Page 53-55; 59pp; English.		
XX	The invention relates to testa specific Cv20ox gene, a gibberellin		
CC	20-oxidase (GA 20-oxidase) gene which is integument-specifically		
CC	expressed in seeds of the Citrullus lanatus, and proteins encoded		
CC	by such genes. The invention also relates to Cv20ox promoter		

CC	CC (cv20oxp)	CC which regulate the expression of cv20ox gene and method for generating seedless fruits of gourd family such as melon or watermelon using the cv20oxp sequence. The present sequence is C. lanatus cv20ox cdna coding for GA 20-oxidase.	xx
SQ	Sequence 1420 BP; 412 A; 257 C; 329 G; 422 T; 0 other;		
Query Match	9.2%; Score 229.6; DB 24; Length 1420;		
Best Local Similarity	55.3%; Pred No. 3.7e-36;		
Matches 501; Conservative	0; Mismatches 384; Indels 21; Gaps		
QY	926	CGCGCAGTGGGGGAGCGTGCAGAGACGACGGCTTCTCTCTGGTGGTTAAACACGGCAT 985	
DB	288	CAGGCTGGTGATGAAGCGTGTAAACATCATGGATCTTTGTTGGTGAACCATGGAGT 347	
QY	986	CGAGCGGGCTGCTGGAGAGCGCACCGGTGCATGAGAGCCCTTTCACGCTGCCTCT 1045	
DB	348	TGATATGGAATGTATAAAGGGTTTCATGAGTGTATGATGAGTTTTTTCACATTGCCCTT 407	
QY	1046	GGGGAGAGACGAGCGCGCAGCGCGCGGGGGAGAGCTGGCGCTACGCCAGCAGCTT 1105	
DB	408	TGATGTGAACAAAATCTCAAAGAAATTTGGTGAATAATGGGTATGCTATAGCTT 467	
QY	1106	CAGCGGCGCTTCGCTCCAAAGTCGCGGTGGAAGAGACGCTGCTGTTCCGGTACTCATC 1165	
DB	468	CATTGGAAGATTCCTCAATAATCTCCATGGAAGAAACCTTTCCCTTC----- 517	
QY	1166	GGTGAGATGAAGAGGGCGAGAGGGCGGTGGGTAGTACTCGTGGCGGAAGCTCGGGG 1225	
DB	518	--CTTAATGCTCTAATCAATAATCCACCTGTTTCAGGATTTGGTTTCAATGTGTGTGG 575	
QY	1226	GGAGCAGCGGGCGCTGGCGAGGTGACTCGCGTACTGCCAGAGATGAGCGGCT 1285	
DB	576	CCAGAAATATCCCATCAAGGAAGGTGTATCAAGATGTGGAAGAAGCATGAGTGATCT 635	
QY	1286	GTGCTGGAGCTGATGAGGTGTCGCGGGAGAGCGCTGGGATCGTCGGAGACCGCGCCA 1345	
DB	636	GAGTTTGAAGATAGTGGAGCTATTGGGGTTGAGCTAGGGTTCCAAAAG-----A 686	
QY	1346	CTACTTCGGCGATTTCTCAGCGCAAGCACTCATATGCGGCTCAACTACTACCGCGC 1405	
DB	687	GAAGTTCAAAAATTTATGAGCAAGATGATCAATATTAAGCTGAATATTATCCCAAC 746	
QY	1406	GTGCCAGAGGCACCTGCACACGCTGGGACACCGGTTCGACCTGGAGACCCACCTCGCTCAC 1465	
DB	747	ATGGGAGAGCCAGCTGACATTTGGAACTGGGCTCATCTGTGATCCAACTCCATTAAC 806	
QY	1466	CATCCTCCACAGGACCACTCGCGGGCGCTGGAGGTGTGGCGGAGAGGGCGGTGCGCGC 1525	
DB	807	AATTCTTCCAAAGACCACTCATAGTGGCGCTTCAAGTCTAAGTGTGATGATGAGTGCATC 866	
QY	1526	CATCGCGCTCGCCCGGGCGCTGCTGTGTAAGCTGGGGACACCTTCATGGCGCTCTC 1585	
DB	867	AATTCTTCCAAAGAGACTCTTTGTATCAATATCGGTGACATTTTATGGCTCTGAC 926	
QY	1586	CAACGCCAGGTACCGCAGCTGCCTGCACCGGCGGTGTCTCAACAGACACGGCGCTTCGCGC 1645	
DB	927	AATGGGTTTACAGAGCTGCTTCCACGAGCTGTAGTGACTGCAAGAAAGTAAGAA 986	
QY	1646	CTGCTGCGCTTCTCTGTCGGGAGATGGAACACGGTGGTGGCGCCCGCGGAGAGCT 1705	
DB	987	ATCAATGGCTCTTCTTATGTCCAGCGGTGACAAAGTGTGTGAGAGCACCGGAGAGT 1046	
QY	1706	GTTCAGACACCAACCCAGGGGTGTACCGGACTTCACGTGGCGGGCGCTCTGGACTT 1765	
DB	1047	GGTGGAGGAAGATCCACCAAGGAAGTTTCCAGATTACATATGGCCAACTGCTCTTGAAC 1106	
QY	1766	CACGAGCGCCACTACAGGCGGACATGCGACGCTTCAGCGCTTCTCGACTGGCTTAA 1825	
DB	1107	TACCAAAAAATTTATAGGCTGATTCGAACACTTTTCAGGCTTTCATACCTTGGCTTGA 1166	
QY	1826	TCATCA 1831	

QY	1108	CGGGCGCTTCGGCTCCAAAGCTGCCGTGGAGGACGACCTCTCGTTCGGGACTCATCGG	116
DB	427	TTGGGAGATTTCGGCTCCAAATCTTCCATGGAGAAACCTTTTCCTTCGCTGTGTGGCTG	488
QY	1168	CTGGAGATGAAGGGGCGAGGAGGGCGTGGTGTAGTACTGTCGGAAGACTCGGGGGCGG	1227
DB	487	CTCAAACTCCTCCGGGCTCATGACTATGTTCTTGACACTTTAGGCCCATCATTTCTCC	546
QY	1228	AGCACGGGGGGGGTGGGGAGGTGTACTGCGGTACTTGCCACGAGATGACGCCCTGT	1287
DB	547	ATCATGGS-----AAGCGCTATCAAGAGTGTGGGATAGCATTTGAACGAGCTTG	594
QY	1288	CGCTGGAGCTGTATGGAGTGTCTCGGGGAGAGCTGGCATCTGCGGAGACCGGCCACT	1347
DB	595	GTACGAAGATTGTGAGACTTTTGGGGCTTAGCCTTGGCATTTCAAGAG-----AAT	645
QY	1348	ACTTCCGGGATTTTCGAGCGCAACGACTCCATCATCGCTCACTACTACTACCGGGT	1407
DB	646	ACTTCAAGAAATTTCTTCGAGGACCAACGATTCAATATTGAGCGTTTAATTACTACCAACAT	705
QY	1408	GCAGAGGCCACTGCACACGCTGGGSCACCGTCCGACTGCGACCTGGACCCACCTCGCTCAACA	1467
DB	706	GGACAAAGCCAGAGGTGTGTGGGAACCTGGCCCTCACACTGATCCACCTCCGTCACAA	765
QY	1468	TCCTCCACAGGACCAACGCTCGGCGCTCGAGGTGTGGGGCGAGGGGGGTGGGGCGCA	1527
DB	766	TCCTTCACCAAGACCTGTCTAGTGGCTTCAAGTGTCTCCATATGATCAATATGATTTCAA	825
QY	1528	TCGCGCTTCGCGCGGGCGCTCGTCAACGCTGGGCGGACACCTTCATGCGGCTCTCCA	1587
DB	826	TTCTTCCAAACCCAGAGCCTTTGTTCATCAACATCGGTGACACTTTCACCTCTCTCACA	885
QY	1588	ACGCCAGGTACCGAGCTGCCTCCGCGCGGCTCGTCAACAGACCGCGCTTCGCGCTG	1647
DB	886	ATGGGATTTACAAGGCTGCATACACCGCGCTGTAGTGAATTCATGATCAAGAAAT	945
QY	1648	CGCTGGCGCTTCCTCTCTCCCGGAGATGACACGCTGTGGCGCGCGCGCGGAGAGCTGG	1707
DB	946	CATTGGCGCTTCTTCTGTGTCATGATGACAAAGTGTGGAGACACCGGAGGATTTGG	1005
QY	1708	TCGAGCACCCACCGAGGCTTACCGGACTTCACGCTGGGGCGCTGCTGGACTTCA	1767
DB	1006	TCGAGAGAGTCCACACGAAAGTATCCAGATTAAATGGCCATGTTCTTGAATGA	1065
QY	1768	CCACGCGCACTACAGGCGGACATGCGCAGCGTTCAGCGCTTCGCGCTTCGCGCTTAA	1825
DB	1066	CCAAAAGCGTTACCGACCTGATTGCAACACTTTGGAAGCCTTCAAAACTTGGGTCA	1123
RESULT 7			
AAZ99483	AAZ99483	standard; cDNA; 1161 BP.	
XX	AAZ99483		
XX	AAZ99483		
XX	03-JUL-2000	(first entry)	
XX	CDNA	encoding a C-20 oxidase polypeptide.	
DE	Gibberellin acid; copalyl diphosphate synthase; 3beta-hydroxylase;		
DE	2-oxidase; phytoene synthase; C-20 oxidase; 2beta,3beta-hydroxylase;		
KW	seed germination; seedling growth; gibberellin biosynthetic pathway;		
KW	transgenic plant; hypocotyl; epicotyl; ss.		
XX	Cucurbita maxima.		
OS			
XX	Key	Location/Qualifiers	
FE	1..1161		
FT	/tag= a		
FT	/product= "C-20 oxidase"		
XX			
XX	WO200009722-A2.		



CC new breeds of plants, for producing a higher harvest of feeding materials  
 CC for animals and in plant regeneration. They are used to control the  
 CC growth, size and shape of plants, e.g. dwarf plants, by inhibiting the  
 CC gene expression. The present sequence encodes a rice gibberellin  
 CC 3beta-hydroxylase enzyme of the invention.  
 XX  
 SQ Sequence 1122 BP; 137 A; 406 C; 410 G; 169 T; 0 other;

Query Match 7.0%; Score 174; DB 22; Length 1122;  
 Best Local Similarity 49.8%; Pred. No. 3.9e-25;  
 Matches 509; Conservative 0; Mismatches 495; Indels 18; Gaps 2;

QY 799 GCGGCGAGCGGAGATCCGCTCGACATTCATATGCGCGCGGAGAGAGCGCGGCTCGG 858  
 DB 47 GGGCGGCGAGCGGGTGGCGGAGACGCGCGTGGCGCGGGGTGGAGCACCACCGGTTG 106  
 QY 859 TGGCGGTGGAGAGTGGAGTGGCGGTGATCGACGTGGGGGCGGCGGAGAGTCTGT 918  
 DB 107 TGGAGCGCGCGCGCGCGCGGAGGACGCGGTGCGGTGGTGGAGCTCAGGCGGCGG 166  
 QY 919 CGGTGTCGGGAGAGCGGTGGAGAGGACGCGGTTCCTGCTGGTGGTTAAC 978  
 DB 167 ACGCGGCGCGCGGTGGCGGCGGCGCGGAGCAGTGGCGGCTCTCTGCTGGTGG 226  
 QY 979 ACGGCTACGAGCGCGCTGTGGAGGAGCGCGACCGGTGATGAGCGCTTTCACGC 1038  
 DB 227 ACGGCTGCGCGCGCGCTGTGTGCGGCTCGAGAGCGCTGCCGCGGTCTTCC 286  
 QY 1039 TGGCGGTGGGAGAGAGAGCGGCGCGGCGCGGAGAGTGGCGGTACGCCA 1098  
 DB 287 TGGCGGCTGGAGAGATGGCGCGCTCGCGCGCGCGCGGAGCGCTTGGCGGTACGG 346  
 QY 1099 GCAGCTTACGCGGCGCTTCGCTCCAGCTGCGGTGGAGAGAGCGTTCGTTCCGT 1158  
 DB 347 CGCGCGCCATCTCTCTCTTCCAGCTCATGTGTTGGTCCGAGGCTTACACCTTCC 406  
 QY 1159 ACTCATCGGTGAGATGAAGAGGCGGAGGCGGTGGGTGAGTACCTGTGGGAAGC 1218  
 DB 407 CTTCCTCCCTCGCTCGCGAGCTCGCGCGCTCTGGCCCAAGTCCGCGGAGACTCTC 466  
 QY 1219 TCGGCGGCGAGCAGCGGCGGCGGTGGCGAGGTGTACTTCGCGTGTACGACGATGA 1278  
 DB 467 TCTTCTGTGACGTGTGAGGAGTTTCCAAAGAGATGGCGGCTAGCGGAGTTCG 526  
 QY 1279 GCGCGCTGTGCTGAGCTGTAGTGTGCTCGGCGAGAGCTGGCATCTCGGAGACC 1338  
 DB 527 TGAGTTGTCTTGTAGGCGCTGGGGTCTACCGCGGAGAGTCCCGGAGTGGGCGG 586  
 QY 1339 GCGCGCACTACTTCCGCGGATTTCTTCCAGCGCAACGACTTCCATCATGCGCTCA 1398  
 DB 587 AG-----AGGAGATCGCGAGAGGATGACGCGAGCTGCACTCACTGTGT 634  
 QY 1399 ACCGCGGTGCGAGAGCGGCTTCCAGAGCTGTGGCGACGCTCGGCTGCGACGACCT 1458  
 DB 635 ACCGAGGTGCGCGGAGCGGCGGAGCGGTGGGGTCTATCGCGACAGGACTTCGGGT 594  
 QY 1459 CGCTTACACATCCCTCCAGGAGACAGCTGCGGCGCTGAGGTGT-----GGCGGAGG 1512  
 DB 695 TCTTACCTCTGTCTCAGAGAGCTCTCGCGGCGTGGAGTGTCTCGTGGAGGCGCG 754  
 QY 1513 GCGGCTGGCGGCGGCTTCCGCGGCTCGCGGCGGCGCTCTGCTCAAGTGGGAGACCT 1572  
 DB 755 ACCGCTGGGTGGGTTGGCGGCGGTGGCGGCGGCTCTGCTCAAGTGGGAGCTCT 814  
 QY 1573 TCATGGGCTCTCCAGCGGAGTACCGGAGCTGCTGCGACCGGCGGTGCTCAGAGCA 1632  
 DB 815 TCCACATCTTCCAGAACGCGGCTTCCACAGCGCTTACACCGCGCGCTGTGAGCGCG 874  
 QY 1633 CGGCGCTCTCGCTGGGCTCTCTCTCTGCGCGGAGATGGACACGCTGGTGGCGC 1692  
 DB 875 ACCGCGAGCGGGTCTGCTGCTGCTACTTCTCTGCGCGCGCGCGCGGAGGTGGCG 934  
 QY 1693 CGCGGAGAGTGTGTGAGACGACACCGCGGAGGTGTACCGGAGTACCGGCTACGTGGCGG 1752

DB 935 CGGTGCGGAGGCGGTGCGCGCGGCGGAGCGCGCTACCGGCTGTCAAGTGGCGG 994  
 QY 1753 CGCTGTGACTTACGAGCGGCGGCTTACAGGCGGAGATCGCAAGCTTCAGGCGTCT 1812  
 DB 995 AGTACATGCGCGTCCGAGAGGCTTCGCCACGCGGCGTCCGCCCTCAAGATGCT 1054  
 QY 1813 CC 1814  
 DB 1055 CC 1056

RESULT 9  
 AAZ99421  
 ID AAZ99421 standard; DNA; 1077 BP.

XX AAZ99421;  
 XX AC  
 XX AAZ99421;  
 XX 03-JUL-2000 (first entry)  
 XX Soybean C-20 oxidase polynucleotide sequence.  
 DE Gibberellic acid; copalyl diphosphate synthase; 3beta-hydroxylase;  
 XX 2-oxidase; phytoene synthase; C-20 oxidase; 2beta,3beta-hydroxylase;  
 KW seed germination; seedling growth; gibberellin biosynthetic pathway;  
 KW transgenic plant; hypocotyl; epicotyl; ss.  
 XX Glycine max.  
 OS  
 XX W0200009722-A2.  
 XX 24-FEB-2000.  
 XX 10-AUG-1999; 99WO-US18086.  
 XX 10-AUG-1998; 98US-0096111.  
 PR 07-JUN-1999; 99US-0137977.  
 XX (MONS ) MONSANTO CO.

XX Brown SM, Elich TD, Heck GR, Kishore GM, Logusch EW, Logusch SJ;  
 PI Piller KM, Rao SJ, Ream JE;  
 XX WPI; 2000-224351/19.  
 DR  
 XX Obtaining transgenic plant useful for controlling seed germination and  
 PT seedling growth comprises transgene comprising a sequence expressing  
 PT altered levels of an essential hormone. -  
 XX Claim 45; Page 227-228; 267pp; English.

CC The present sequence represents the C-20 oxidase polynucleotide  
 CC sequence. The polynucleotide sequence is used in the method of the  
 CC invention. The specification describes methods for the inhibition and  
 CC control of gibberellic acid levels. Gibberellic acid levels may be  
 CC inhibited or controlled by use of a chimeric expression construct  
 CC expressing a RNA or protein which suppresses the gibberellin  
 CC biosynthetic pathway sequence, diversifies substrate from the pathway, or  
 CC degrades pathway substrates or products. The methods uses copalyl  
 CC diphosphate synthase, 3beta-hydroxylase, 2-oxidase, phytoene synthase,  
 CC C-20 oxidase, and a 2beta,3beta-hydroxylase polynucleotides to achieve  
 CC this. The method is used to control seed germination and seedling  
 CC growth especially to regulate gene products of gibberellin biosynthetic  
 CC pathway and restoration of normal seed germination, in transgenic  
 CC plants. The plants produced are gibberellin deficient, and have  
 CC shortened hypocotyl and/or epicotyl phenotypes compared to normal  
 CC plants.

XX SQ Sequence 1077 BP; 316 A; 219 C; 248 G; 288 T; 6 other;

Query Match 6.1%; Score 153.2; DB 21; Length 1077;  
 Best Local Similarity 50.1%; Pred. No. 5.2e-21;  
 Matches 470; Conservative 1; Mismatches 450; Indels 18; Gaps 3;

XX	Maize gibberellic acid 2-oxidase 2 cDNA sequence.
DE	
XX	Gibberellic acid; copalyl diphosphate synthase; 3beta-hydroxylase;
DE	2-oxidase; phytoene synthase; C-20 oxidase; 2beta,3beta-hydroxylase;
KW	seed germination; seedling growth; gibberellin biosynthetic pathway;
KW	transgenic plant; hypocotyl; epicotyl; ss.
KW	
OS	Zea mays.
XX	
PN	WO200009722-A2.
PN	
PD	24-FEB-2000.
XX	
PF	10-AUG-1999; 99WO-US18066.
XX	
PR	10-AUG-1998; 98US-0096111.
PR	07-JUN-1999; 99US-0137977.
XX	
PA	(MONS ) MONSANTO CO.
PI	Brown SM, Elich TD, Heck GR, Kishore GM, Logusch EW, Logusch SJ;
PI	Piller KJ, Rao S, Ream JE;
DR	WPI; 2000-224351/19.
XX	
PT	Obtaining transgenic plant useful for controlling seed germination and
PT	seedling growth comprises transgene comprising a sequence expressing
PT	altered levels of an essential hormone -
XX	
PS	Claim 45; Page 253; 267pp; English.
XX	
CC	The present sequence represent a gibberellic acid 2-oxidase 2 cDNA
CC	sequence, which is used in the method of the invention. The specificat
CC	describes methods for the inhibition and control of gibberellic acid
CC	levels. Gibberellic acid levels may be inhibited or controlled by use
CC	a chimeric expression construct expressing a RNA or protein which
CC	suppresses the gibberellin biosynthetic pathway sequence, diverts
CC	substrate from the pathway, or degrades pathway substrates or products
CC	The methods uses copalyl diphosphate synthase, 3beta-hydroxylase,
CC	2-oxidase, phytoene synthase, C-20 oxidase, and a 2beta,3beta-hydroxyl
CC	polynucleotides to achieve this. The method is used to control seed
CC	germination and seedling growth especially to regulate gene products of
CC	gibberellin biosynthetic pathway and restoration of normal seed
CC	germination, in transgenic plants. The plants produced are gibberellin
CC	deficient, and have shortened hypocotyl and/or epicotyl phenotypes
CC	compared to normal plants.
XX	
SQ	Sequence 811 BP; 134 A; 249 C; 278 G; 150 T; 0 other;

Query Match	5.2%;	Score 129;	DB 21;	Length 811;
Best Local Similarity	56.9%;	Pred. No. 3.1e-16;		
Matches 278;	Conservative 0;	Mismatches 205;	Indels 6;	Gaps 2;
QY 1274	GATGAGCGCGCTGTGCTGGACCTGATGGAGGTGCTWCGGGGAGAGCCTGGGCGATCGTCGG 1333			
Db 22	GATGGCTGTGCGCGTCTGTGGAGCTGATGGCGGAGGGGCGTGGGCATCGCGCGCGCGCGCG 81			
QY 1334	AGAGCGGGCGCACTACTTCCGCGGATCTTCCAGCGCAACAGACTCCATCATGCGGCTCAA 1393			
Db 82	GGACCGGTGCTGGCGCGCTGTGTGGCGCGCGCGGACACGACTGCTGATGTGTGGGTGAA 141			
QY 1394	CTACTACCGCGCGTGGCAGAGCCACTGCACGCT---GGGCAACCGGTCGCCACATGCGA 1450			
Db 142	CCACTACCGCGCGCGCGCTCAACCCAGCCTACGGGGTTCGGCGAGCACACCGA 201			
QY 1451	CCCACTTGCTGCATCCTCTCCACGAGACACAGCTGGCGGCGCTGGAG---GTGTGGGC 1507			
Db 202	CCCGAGATCATCTGCTGCTCCGGCGCAACGGGCACCTCCGCGCTGGAGATCGCGTGG 261			
QY 1508	GGAGGGCGGTGGCGCGCATCCGCCCTCGCCCCCGGGGCGCTGCTGCTCAACGTCCGGCA 1567			
Db 262	GGAGGGCGCTTGGCGCTCCGTCGCCGCCCGACGGGGAGCGCTTCTTCGTCAACGTCCGGCA 321			

RESULT 10  
AAZ99478  
ID AAZ99478 standard; CDNA; 811 BP.  
XX  
XX AAZ99478;  
XX AC  
DT 03-JUL-2000 (first entry)







CC encoding the stabilising polypeptide can be linked onto or inserted into  
CC a nucleic acid encoding a core protein. The fusion proteins of the  
CC invention are more resistant to degradation by proteases and, thus, have  
CC a longer half-life than the unfused core protein. The products can be  
CC used for treating autoimmune diseases, cancer and inflammation. In  
CC particular, the core protein may be an IkappaB regulator protein for the  
CC treatment of inflammatory bowel disease, or a nitroreductase protein  
CC which can activate nitro drugs in enzyme/prodrug therapy to treat cancer  
CC or other pathological conditions. The fusion proteins can also be used in  
CC diagnostic methods such as in vivo imaging.

XX SQ Sequence 799 BP; 201 A; 106 C; 479 G; 13 T; 0 other;

Query Match 3.8%; Score 94.4; DB 19; Length 799;  
Best Local Similarity 49.2%; Pred. No. 2.3e-09;  
Matches 248; Conservative 0; Mismatches 256; Indels 0; Gaps 0;

QY 752 GGTGGTGCACGAGGAGGCGTGTCTGACGCGGCGGTGTGAGCGGGCAGACGCA 811  
DB 68 GGACGAGGAGGGGCGACGAGGAGGAGGGGCGACGAGGAGGAGGAGGAGG 127  
QY 812 GATCCCTCCAGTTCATATGCGCGGCGAGGAGAGCCCGGGTCCGTGGAGGA 871  
DB 128 GCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 187  
QY 872 CTTGAGGTCGCTGATCGACGTGGGGCGGGCGGAGAGGTCGTCGTTGCCGCA 931  
DB 188 CGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 247  
QY 932 GGTGGGAGGCGTGCAGAGAGCGCGCTTCTTCCTGTGGTTCACACGATCGAGG 991  
DB 248 GGAGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 307  
QY 992 GCGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1051  
DB 308 CGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 367  
QY 1052 GAAGCAGCGGCGCAGCGCGGCGGAGAGCTCGCGCTACGCCAGCTTCACGG 1111  
DB 368 SCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 427  
QY 1112 GCGCTTCGCTCCAGCTGCGCGGAGGAGAGCGTGTCTGTTCCGTTACCTG 1171  
DB 428 CGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 487  
QY 1172 AGATGAGAGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1231  
DB 488 GGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 547  
QY 1232 CGGCGGCGGCTGGGCGAGGTGTA 1255  
DB 548 GGAGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 571

RESULT 14

AAA50254  
ID AAA50254 standard; DNA; 1926 BP.

XX AC AAA50254;

XX DT 07-NOV-2000 (first entry)

XX DE Epstein Barr virus nuclear antigen 1 protein (EBNA1) DNA.

XX KW EBV; nuclear antigen 1; EBNA1; episome; transfection; selection;

XX KW gene therapy; ds.

XX OS Epstein-barr virus.

XX PN W020004778-A1.

XX PD 17-AUG-2000.

PF 11-FEB-2000; 2000WO-US03547.  
XX PR 11-FEB-1999; 99US-0249585.  
XX PA (PHAR-) PHARMACOPETIA INC.  
XX FI Horlick RA, Chelsky D;  
XX DR WPI; 2000-515062/46.  
XX DR P-PSDB; AA95856.  
XX PT Stably transfecting eukaryotic cells with at least one episome for the  
XX production of a desired protein in vitro and for gene therapy -  
XX Disclosure; Fig 2: 53pp; English.

CC The present sequence is that of DNA encoding the Epstein-Barr virus  
CC (EBV) nuclear antigen protein 1 (EBNA1, see AA95856). EBNA1 is  
CC utilised in a novel method for obtaining a eukaryotic cell that is  
CC stably transfected with at least one episome. This method involves  
CC transfecting a eukaryotic cell with: (1) a first episome comprising  
CC an EBV origin of replication (oriP, see AAA50253), a gene encoding a  
CC first protein whose expression results in cell death and a  
CC selectable marker for eukaryotic cells; and (2) a second episome  
CC comprising an EBV oriP and a gene encoding a second protein, where  
CC expression of the second protein prohibits the occurrence of cell  
CC death resulting from expression of the first protein to produce  
CC doubly transfected cells which also express an antigen that  
CC promotes retention of the episomes by the cells. The doubly  
CC transfected cells are maintained under conditions in which the  
CC first and second proteins and the selectable marker are expressed,  
CC and the selective pressure specified by the marker is maintained.  
CC Under these conditions, only cells containing both episomes live.  
CC Preferably, EBNA1 is expressed from 1 of the episomes, and the  
CC protein of interest from the other episome. Either or both  
CC epitopes may further comprise a nucleic acid sequence encoding a  
CC protein, a nucleic acid encoding an RNA that is not intended to  
CC be translated (e.g. a therapeutic RNA), or a DNA sequence used as  
CC a tag for the cells. The method is applicable to cell culture or  
CC intact organisms, for gene therapy. It allows the rapid  
CC establishment of eukaryotic cells that stably and reliably express  
CC a gene of interest, using a novel method of selection, and  
CC maintenance of that selection without the need for exogenous  
CC selection factors, such as antibiotics.

XX SQ Sequence 1926 BP; 487 A; 352 C; 872 G; 215 T; 0 other;

Query Match 3.8%; Score 94.4; DB 21; Length 1926;  
Best Local Similarity 49.2%; Pred. No. 2.9e-09;  
Matches 248; Conservative 0; Mismatches 256; Indels 0; Gaps 0;

QY 752 GGTGGTGCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 811  
DB 310 GGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 369  
QY 812 GATCCCTCCAGTTCATATGCGCGGCGAGGAGAGCCCGGGTCCGTGGAGGA 871  
DB 370 GCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 429  
QY 872 GCTGAGAGTGGCGCTGATCGACGTGGGGCGGGCGGAGAGGTCGTCGTTGCCGCA 931  
DB 430 GCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 489  
QY 932 GGTGGGGGAGGCGTCGAGAGGACGCGCTTCTTCTGTGGTTCACACGATCGAGG 991  
DB 490 GGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 549  
QY 992 GGGCGCTCTGGGAGGAGGCGCACCGGTGATGAGCGCTTCTTCAGCTGCCGTGGGGA 1051  
DB 550 GCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 609  
QY 1052 GAAGCAGCGGCGCAGCGCGCGCGGAGAGAGCTCGCGCTACGCCAGCTTCACGG 1111



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2003, 19:50:00 ; Search time 2548.43 Seconds  
(without alignments)  
15894.084 Million cell updates/sec

Title: US-09-702-134-7212\_COPY\_28000\_30500

Perfect score: 2501

Sequence: 1 atgcgaagataaaattggc.....ctatcaatggttgcttcaat 2501

Scoring table: IDENTITY\_NUC  
Gap 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estlin:\*
- 4: em\_estma:\*
- 5: em\_estov:\*
- 6: em\_estpi:\*
- 7: em\_estro:\*
- 8: em\_htc:\*
- 9: gb\_esti:\*
- 10: gb\_est2:\*
- 11: gb\_htc:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estfun:\*
- 16: em\_estom:\*
- 17: gb\_gss:\*
- 18: em\_gss\_hum:\*
- 19: em\_gss\_inv:\*
- 20: em\_gss\_pln:\*
- 21: em\_gss\_vrt:\*
- 22: em\_gss\_fun:\*
- 23: em\_gss\_mam:\*
- 24: em\_gss\_mus:\*
- 25: em\_gss\_other:\*
- 26: em\_gss\_pro:\*
- 27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	629.4	25.2	1435	11	AY105651 Zea mays
2	286.8	11.5	536	12	BG322884 EML15.C1
3	277.4	11.1	755	12	BG365195 HVSMH000
4	263.4	10.5	752	12	BG365194 HVSMH000
5	261.4	10.5	796	12	BG299451 HVSMH002
6	255.4	10.2	577	10	BE194605 HVSMH008

7	230.4	9.2	914	12	BF620505
8	213	8.5	682	12	BG416635
9	209.8	8.4	566	12	BG414208
10	203.8	8.1	571	10	BG454826
11	200	8.0	567	14	BG065653
12	196.8	7.9	655	12	BG350197
13	191.2	7.6	517	10	BG422495
14	187	7.5	686	10	BG414214
15	185.6	7.4	700	14	BG402912
16	173.6	6.9	757	10	BG055752
17	170.8	6.8	588	10	BG429459
18	167.8	6.7	483	13	BF950763
19	158.6	6.3	801	12	BG369047
20	156	6.2	1084	11	AY108935
21	151	6.0	612	14	BG086836
22	142.8	5.7	472	17	AG914586
23	142.2	5.7	438	14	BG167179
24	141.6	5.7	564	10	AW471536
25	140.6	5.6	383	13	BF131287
26	140	5.6	350	10	BE500119
27	138.6	5.5	513	9	AL506369
28	135.4	5.4	773	14	BG743979
29	134.8	5.4	405	10	BE556197
30	134.4	5.4	675	14	BG402913
31	133.6	5.3	590	10	AV563665
32	133.6	5.3	662	12	BE918296
33	130.6	5.2	412	12	BF293852
34	130.6	5.2	583	10	AV560858
35	130.4	5.2	521	12	BE918082
36	130.2	5.2	761	14	BG0807183
37	130	5.2	666	14	BG413073
38	129.8	5.2	803	10	AW448262
39	129.8	5.2	803	14	BG605936
40	128.6	5.1	1382	11	AY104706
41	128.4	5.1	364	13	BF124052
42	128.4	5.1	628	9	AL504453
43	127	5.1	665	14	BG415927
44	126.6	5.1	472	10	BG403224
45	126.4	5.1	479	13	BM136125

#### ALIGNMENTS

RESULT 1	AY105651	AY105651	1435 bp	mRNA	linear	HTC 25-MAY-2002
LOCUS	Zea mays PC0130567 mRNA sequence.					
DEFINITION	Zea mays PC0130567 mRNA sequence.					
ACCESSION	AY105651					
VERSION	AY105651.1					GI:21208729
KEYWORDS	HTC.					
SOURCE	Zea mays.					
ORGANISM	Zea mays					
REFERENCE	1 (bases 1 to 1435)					
AUTHORS	Bailey C.F., Dolan M., Miao G.H., Vogel J.M., Whittitt, M.S., Arthur D.W., Hanarey, M., Morgante, M. and Tingey S.V. Design of Maize Mapping Project/DuPont Consensus Sequences for Overgo Probes Unpublished (2002)					
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Overgo Probes					
JOURNAL	Unpublished (2002)					
AUTHORS	Coe, E.C.					
TITLE	Direct Submission					
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA					
FEATURES	Location/Qualifiers					
SOURCE	1..1435					
	/organism="Zea mays"					
	/db_xref="MaizeDB:637612"					
	/db_xref="taxon:4577"					
	/clone="PC0130567"					

/clone\_lib="Maize Mapping Project/DuPont Cornsensus Library"  
/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed Dupont contigs; this resource was assembled by Dupont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"  
BASE COUNT 284 a 461 c 427 g 263 t  
ORIGIN

Query Match 25.2%; Score 629.4; DB 11; Length 1435;  
Best Local Similarity 75.9%; Pred. NO. 1.1e-118;  
Matches 825; Conservative 0; Mismatches 241; Indels 21; Gaps 3;  
  
QY 744 ATGAGCATGGTGGTGCAGCAGGAGGAGGTTCAGACGGCGGTGCTGAGCGGG 803  
DB 67 ATGGTCGTGGTGGCGACAGTCCCTCCCTTGTTCGACGCTGCCCGCTGAGCGGC 126  
  
QY 804 CAGACGAGATCCCGTCGCATTCATATGSCCGCGGAGAGAGACCCCGGTCGGTGGCG 863  
DB 127 CTCTCCGACATCCCGACAGTTCATCTGCGCGGCGAGAGAGCCACCCCGGACTCC 186  
  
QY 864 GTGGAGGACTGGAGGTGGCGCTGATCAGAGTGGGGGGGGGGGGAGAGGTGCTCGCGT 923  
DB 187 GCGAGGAGCTGGCGCTGCGCGCTCATCGACCTCTCCGGGGAGCGCG-----CCGAGGTG 240  
  
QY 924 GTCCGCAAGTGGGGAGGCTGCGAGAGGACGCGCTTCTTCTTCTTCAACACCGC 983  
DB 241 GTCCGCAAGTGGGGAGGCTGCGAGGCTGCGAGTTCACGCTTCTTCAAGTGGTGGGACCGC 300  
  
QY 984 ATCGAGCGGCGCTGTGGAGAGGCGCACCGGTGATGAGAGCGCTTCTTCAAGTGGCG 1043  
DB 301 ATCGAGCGGCGCTGTGGAGAGGCGCACCGGTGATGAGAGCGCTTCTTCAAGTGGCG 360  
  
QY 1044 CTGGGGAGAGCAGCGCGCGCGCGCGGAGAGCTGCGCTACGCCAGCGC 1103  
DB 361 CTCCGCAAGTGGGGAGGCTGCGAGGCTGCGAGTTCACGCTTCTTCAAGTGGTGGGACCGC 420  
  
QY 1104 TTCAGCGGCGCTTCCGCTCCAGCTGCGTGGAGAGAGAGCTGCTTCCGCTACTCA 1163  
DB 421 TTCAGCGGCGCTTCCGCTCCAGCTGCGTGGAGAGAGAGCTGCTTCCGCTACTCA 480  
  
QY 1164 TCGCTGAGATGAGAGGCGAGAGGCGTGGGTGATGATGATGATGATGATGATGATG 1223  
DB 481 GAGCAGCAGCAGCGCGACAGTCCAGAGAGCTGCTGGCTTCTTCTTCTTCTTCTTCT 540  
  
QY 1224 GCGGAGCAGCGCGCGCTGGCGAGGTGTACTCGCGCTACTGCCACGAGATGAGCGC 1283  
DB 541 GCGGAGGCTACCGGACCGAGGAGGTGTACTCGCGCTACTGCCACGAGATGAGCGC 600  
  
QY 1284 CTGTGCTGAGAGCTGATGGAGGTGCTCGGGGAGAGCTTGGGCAFTGCTGGAGAGCGGCGC 1343  
DB 601 CTGTGCTGAGAGCTGATGGAGGTGCTCGGGGAGAGCTTGGGCAFTGCTGGAGAGCGGCGC 651  
  
QY 1344 CACTACTTCCGCGATTCCTCCAGCGACAGCTACCATGATGATGATGATGATGATGATG 1403  
DB 652 CGCCTACTTCCGCGATTCCTCCAGCGACAGCTACCATGATGATGATGATGATGATGATG 711  
  
QY 1404 GGTGTCAGAGCGCACTCGACAGCTGGGCGACCGGTGCGGCTGCGGCTGCGGCTGCGT 1463  
DB 712 CGGTGCGAGCGGCTTACGACAGCTGGGCGGCGGCTGCGGCTGCGGCTGCGGCTGCGT 771  
  
QY 1464 ACCATCTCCAGAGGACCGCTGCGGCGGCTGAGAGGTG-----GGGCGGAGGGGCGG 1517  
DB 772 ACCATCTCCAGAGGACCGCTGCGGCGGCTGAGAGGTG-----GGGCGGAGGGGCGG 831  
  
QY 1518 TGGCGGCGCATCCGCGCTTCCGCGGCGGCTGCTGCTCAACGCTGCGGACACCTTCATG 1577  
DB 832 TGGCGGCGCATCCGCGCTTCCGCGGCGGCTTCCGCGGCGGCTTCCGCGGCGGCTTCCGCGG 891  
  
QY 1578 GCCTCTCCAGCGGAGTACCGAGCTGCCTGACCGGCGGCTGCTGCTCAACAGCAGCGG 1637  
DB 932 GCCTCTCCAGCGGAGTACCGAGCTGCCTGACCGGCGGCTGCTGCTCAACAGCAGCGG 122

DB 892 GCCTCTCCAGCGGCGCTAGGAGCTGCTCCACCGCGCGCTCCTCAACAGCGGGTG 951  
QY 1638 CTTCCGCGCTGCTGGCGTCTTCCCTTCCCGGAGATGGACAGGCTGGTGGCGCGCG 1697  
DB 952 GCAGCGCGCTGCTGGCGTCTTCCCTTCCCGGAGATGGACAGGCTGGTGGCGCGCG 1011  
QY 1698 GAGGAGCTGGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCG 1757  
DB 1012 AAGGAGCTGGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCG 1071  
  
QY 1758 CTGGAGCTTCAAGCGCGCGCTTACAGGCGCGAGATGCGACGCTTCAAGGCTTCTCCGAG 1817  
DB 1072 CTGGAGCTTCAAGCGCGCGCTTACAGGCGCGAGATGCGACGCTTCAAGGCTTCTCCAC 1131  
  
QY 1818 TGGCTTA 1824  
DB 1132 TGGCTCA 1138  
  
RESULT 2  
LOCUS BG322884 536 bp mRNA linear EST 27-FEB-2001  
DEFINITION EML\_15\_C10.b1\_A002 Embryo 1 (EML) Sorghum bicolor cDNA, mRNA  
sequence.  
ACCESSION BG322884  
VERSION BG322884.1 GI:13152562  
KEYWORDS EST.  
SOURCE sorghum.  
ORGANISM Sorghum bicolor  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum.  
REFERENCE 1 (bases 1 to 536)  
Reid,S.P., Cordonnier-Pratt,M.-M., Gingle,A. and Pratt,L.H.  
An EST database from Sorghum: developing embryos  
Contact: Cordonnier-Pratt MM  
Unpublished (2000)  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu  
Sequences have been trimmed to exclude polyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.  
Seq primer: JEN REV  
High quality sequence stop: 508  
POLYA=No.  
FEATURES  
source  
1..536  
/organism="Sorghum bicolor"  
/db\_xref="taxon:4558"  
/clone\_lib="Embryo 1 (EML)"  
/note="Organ: Embryos germinated for 24 hr; Vector: pBluescript II from Lambda Zap II; Site\_1: XhoI; Site\_2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision."  
BASE COUNT 85 a 204 c 164 g 83 t  
ORIGIN  
  
Query Match 11.5%; Score 286.8; DB 12; Length 536;  
Best Local Similarity 76.6%; Pred. NO. 1.8e-48;  
Matches 418; Conservative 0; Mismatches 92; Indels 36; Gaps 4;  
  
QY 1098 ACCAGCTTCAAGGCGCTTCCGCTTCCAGCTGCGTGGAGAGGAGCGCTGCTGCTCGG 1157  
DB 9 ACCAGCTTCAAGGCGCTTCCGCTTCCAGCTGCGTGGAGAGGAGCGCTGCTGCTCGG 68  
  
QY 1158 TACTCATCGCTGGAGATGAGAGGCGGAGGCGGTGGGTGAGTACCTGCTGCGGAG 1217  
DB 69 TACT-----CGACGACGAGGCGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 122

```

QY 1218 CTGGGGGCGGAGCAGGGCGGCGCTGGGGGGGGTGTACTCGCGCTACTCCACGAGATG 1277
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 123 CTGGGGGATGCGTAC--CGTCACCGGGAGGGTGTACGGCCGCTACTCTCCGAGATG 179
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1278 AGCCGGCTGTGCTGGAGCTGATGGAGTGTCTCGGGAGAGCCCTGGCATCTGTCGGAGAC 1337
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 180 AGCCGGCTGTGCTGGAGCTGATGGAGTGTCTCGGGAGAGCCCTGGCATCTGTCGGAGAC 231
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1338 CGGGGCACTACTTCGGCGGATTCCTCAGCGCAACGACTCCATCATGCGCCCTAACTAC 1397
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 232 -GCGCGCCACTTCTCCGCGCTTCTCCAGGGGAACGACTCCATCATGCGCCCTAACTAC 290
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1398 TACCGGGGTGCGCAGAGCCACTCGACAGCTGGCAGCGGTTCGCGACTCGACCCACC 1457
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 291 TACCGGGGTGCGCAGAGCCACTCGACAGCTGGCAGCGGTTCGCGACTCGACCCACC 350
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1458 TCCTCACCATCTCCACAGGACCACTGTCGGCGGCTCGGAGGTGTGGGGG----- 1509
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 351 TCCTCACCATCTCCACAGGACGAGCTCGGAGGCTTCGAGGTGTCGACGCCGCCACC 410
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1510 -----AGGGCGGTGGCGGCACTCGCGCTTCGCGCCCGGGGGCGCTGTCGTCAC 1559
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 411 GCGCGCGCAGCGCGCTGCGCTCATCAGGCGCCACCGCGCGCTTCGTCGTCAC 470
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1560 GTGGCGCACCTTCATGGGCTCTCCACAGCGAGTACCGACTCGCTGCACGGGGG 1619
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 471 ATCGGCGACCTTCATGGGCTCTCCACAGCGAGTACAGGAGCTGCTGCGACCGCGC 530
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1620 GTGTC 1625
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 531 GTGTC 536
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
LOCUS BG365195 755 bp mRNA linear EST 22-OCT-2001
DEFINITION HVSMEI0001L04f Hordeum vulgare 20 DAP spike EST library HVCDNA0010
(20 DAP) Hordeum vulgare cDNA clone HVSMEI0001L04f, mRNA sequence.
ACCESSION BG365195
VERSION BG365195.2 GI:163233591
KEYWORDS EST.
SOURCE Hordeum vulgare
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
Poaceae; Hordeum.
1 (bases 1 to 755)
Wing, R., Close, T.J., Kleinbols, A., Wise, R., Begum, D., Frisch, D., Yu
, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi, D.W., Fenton
, R.D., Close, S.J., Oates, R. and Main, D.
Development of a genetically and physically anchored EST resource
for barley genomics: Morex 20 DAP spike cDNA library
Unpublished (2001)
On Mar 8, 2001 this sequence version replaced gi:13254292.
Contact: wing ra
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hq bases = 396
Seq primer: AATTAACCTTCACTAAGGG
High quality sequence stop: 623.
Location/Qualifiers
1..755
/organism="Hordeum vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"
/clone="HVSMEI0001L04f"
/clone_lib="Hordeum vulgare 20 DAP spike EST library
HVCDNA0010 (20 DAP)"

```

/tissue\_type="20 DAP spike"

/lab\_host="SOLR"

/note="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI;  
Plants were grown in the greenhouse at the University of  
California, Riverside (Penton, SJ Close, TJ Close). Whole  
spikes with awns trimmed were collected at 20 DAP (Penton  
) . Total RNA was prepared, poly(A) RNA was purified, one  
pfu were in vivo excised to give pBluescript SK(-) cDNA  
phagemids in the TJ Close lab at the University of  
California, Riverside (Choi). Phagemids were plated and  
picked at the Clemson University Genomics Institute (CUGI)  
(Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA  
preparations, DNA sequencing and sequence analysis were  
performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates  
, Rambo, Main). The sequence has been trimmed to remove  
vector sequence and contains a minimum of 100 bases of  
phred value 20 or above. For more details on library  
preparation and sequence analysis see  
http://www.genome.clemson.edu/projects/barley. To order  
this clone see http://www.genome.clemson.edu/orders Also  
see Close TJ, Wing R, Kleinbols A, Wise R (2001)  
Genetically and physically anchored EST resources for  
barley genomics. Barley Genetics Newsletter 31:29-30.  
(http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)."

BASE COUNT 121 a 256 c 258 g 120 t  
ORIGIN

Query Match 11.1%; Score 277.4; DB 12; Length 755;

Best Local Similarity 65.4%; Pred. No. 1.4e-46;

Matches 477; Conservative 0; Mismatches 231; Indels 21; Gaps 4;

```

QY 797 GAGCGGGAGACGAGATCCCGTGCAGTTCATATGCGCGGAGGAGAGACCCCGGGTC 856
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 GAGCGGAAGGAGGAGCTCCCGCGCGAGTTCACTTGGCCAGCGACGCGCCGTCCT 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 857 GGTGGCGGTGGAGAGCTGGAGTGGCGCTGATCGA---CGTGGGGGGGGGGGAGAG 913
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 GGACGCGGTGGAGGAGATCGCTGCGCGGTTCGCGACCTCCCGCGGTTCCTGCGCGGGA 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 914 GTGCTGCTGCTGCGGACAGGTGGGGAGGCGTGGAGAGGACGCGCTTCTTCTGCTGT 973
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 CGACGCTGCGCTCAATGAGCTGCTGCGCGGTGGAGAGGACGCGGTCTTCTCAGGTCT 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 974 TAACACAGGCACTGAGCGCGCTGTGGAGAGGCGCACCGCGTGCATGAGACGCTTCTT 1033
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 GAACACAGCGGTGGAGCGCGCTGTGCGCAAGCGCTACCGCTGCTGCCACGCTCTTA 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1034 CAGCTGCGCGCTGGGGAGAGAGCGGGCGCGAGCGCGCGGGGGAGAGCTGCGGCTA 1093
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 CGCGCTGCGCTGCGCGAGAGAGCGCGCGCGCGCGCGCGCTTCGGCGAGAACCCACGCTA 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1094 CGCAGCAGCTTTCAGCGGCGCTTTCGCTTCCAGGTGCGTGGAGAGAGAGCGTGTCTGTT 1153
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 CGCGCGACGCTTCTGCGGCGGTTCGCGACAGCTCCCGCTGGAGAGAGACCATGCTCTT 360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1154 CCGGTACTATCGGCTGAGATGAAGAGGCGGAGAGGCGGTGGGTGAGTACCTGTGCG 1213
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 CAACCTGCTCCGCGCGCGGAGAGCGCGCGCGCGCGCGCTGCTGCGCTTCTGCGG 414
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1214 GAAGCTCGGGGGAGCAGCGCGCGGTGGCGAGGTGCTGCGGCTACTGCGACGA 1273
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 415 CGTCTTCGCGGAGGAGTAC---CGTCACATGGGGATGTGTGGCACGAATACTGCAAGA 471
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1274 GATGACCGCTGCTGCTGGAGTGTGAGGTGCTCGGGGAGAGAGCTGGGCATCTGCGG 1333
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 472 GATGACACGCTTGGACACTGAGAGCTCAGGAGGTCTTGGCGCGGTGCTGGGCTCTACGG 531
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1334 AGACGCGCGCACTACTTCCGGCGATCTTCCAGCGCAACGACTCCATCATGCGCTCAA 1393
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 532 AGCGGCGC-----TGGCGGCTTCTTTGCGCGGAGCGAGCTCCCTGATGGGCTGAA 582
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1394 CTACTACCGCGGTGCCAGAGGCGCACTGCGACGCTGGGGACCGGTCGCGACTGCGACCC 1453
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```











Email: rwing@clemson.edu  
total hq bases = 618  
Seq primer: AATAACCTCATAAGGCG  
High quality sequence step: 669.  
Location/Qualifiers

```
1. 682
/organism="Hordeum vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"
/clone="HVSMEK0013M18"
/hvcn_lib="Hordeum vulgare t
HVCN0A0013 {ncrna}"
/tissue_type="testa/pericarp"
/lab_host="TJC121"
/note="vector: lambdaZAP; Site
Plants were raised from seeds
growth chamber maintained in
testa and pericarp were dissected
at Washington State Universit
von Wettstein). Total RNA was
purified, one cDNA library was
were in vivo excised to give
phagemids in the T7 Cloze labo
California, Riverside (Akhuho
Kianian, Otto, Simons, Zhang)
picked at the Clemson Universi
(Begum, Palmer, Frisch, Atkin
preparations, DNA sequencing
performed at CUGI (Wing, Yu,
, Rambo, Main). The sequence
vector sequence and contains
phred value 20 or above. For
preparation and sequence anal
http://www.genome.clemson.edu
this clone see http://www.gen
see Close TJ, Wing R, Kleinbo
Genetically and physically ana
Barley genomics. Barley genet
(http://wheat.pw.usda.gov/ggdp
BASE COUNT      154 a   205 c   184 g   138 t
ORIGIN
Query Match           8 5%    Score 213; DB 1
Best Local Similarity 70.8%;  Prod. NO. 2.2e-3-3
Matches 298; Conservative 0; Mismatches 1
```

QY	1411	AGAGGCCACTGCACACGGTGGGCACCGGTCCGCACATCGCAG
Db	2	AGAAGCCTCACTGACGCTGGGCACGGCCGCCACACCAGCAG
QY	1471	TCCACGAGCACAAGTCGGCGGCTTGAGGTGTGGGGGGA
Db	62	TCCACGAGCACAAGTCGGCGGCTTGAGGTGTTCACCGG
QY	1531	GCCCTCGCCCCGGGGCGCTCGTCTCAACTCGTGGCGACAC
Db	122	GGCCCCGAGCAGACGCCTTCGTCTCACATCGCGACAC
QY	1591	CCAGGTACCGCAGCTTGCCTGCACCGGGCGGTCTGCACAC
Db	182	GGCGCCAGCTCAGCTGCCTCCACCGCGCCCTGCTCAACGCG
QY	1651	TGGCCCTCTTCTCTGCCCAGAGATGACACGGTGTGTCG
Db	242	TCACTTCTTCTCAACCCGAGCTGACCGCCCGCTFCA
QY	1711	AC---GACCACACCCGAGGTTGACCCGACCTTCAGTGT
Db	302	CCATTGATGGACGCCGCGGTGTACCCGACCTTCACTT
QY	1768	CGCAGGCCACTACAGGCGCACATGCACGCTTCAGG
Db	362	CSCAAGCATTACCGTCTCATCTCAGTCGAGGACCTGATG

```

QY 1828 A 1828
Db 422 A 422

RESULT 9
BG414208
LOCUS
DEFINITION
HVSMEH0001C12f Hordeum vulgare testa/pericarp EST library
HVCNDA0013 (normal) Hordeum vulgare cDNA clone HVSMEH0001C12f, mRNA
sequence.
ACCESSION
BG414208
VERSION
BG414208.1 GI:13319759
KEYWORDS
EST.
SOURCE
Hordeum vulgare.
ORGANISM
Hordeum vulgare.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.
1 (bases 1 to 566)
Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Kannangara, G., von
Weinstein, D., Akhunov, E., Chin, A., Choi, D.W., Fenton, R.D., Kianian
, P., Otto, C., Simons, K., Zhang, D., Begum, D., Frisch, D., Yu, Y.,
Henry, D., Palmer, M., Rambo, T., Simmons, J., Oates, R. and Main, D.
Development of a genetically and physically anchored EST resource
for barley genomics: Morex testa/pericarp cDNA library
Unpublished (2001)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hg bases = 168
Seq primer: AATTAACCTCACTAAAGGG
High quality sequence start: 9
High quality sequence stop: 545.
Location/Qualifiers
1..566
/organism="Hordeum vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"
/clone="HVSMEH0001C12f"
/clone.lib="Hordeum vulgare testa/pericarp EST library
HVCNDA0013 (normal)"
/tissue_type="testa/pericarp"
/lab_host="TJC121"
/lab_vector="lambdaZAP; Site_1: EcoRI; Site_2: XhoI;
Notes: Vectors were raised from seeds in a Controlled Environments
growth chamber maintained in continuous light at 18°C, and
testa and pericarp were dissected from developing kernels
at Washington State University, Pullman, WA (Kannangara,
von Wettstein). Total RNA was prepared, poly(A) RNA was
purified, one cDNA library was made, and 1 million pfu
were in vivo excised to give pBluescript SK(-) cDNA
phagemids in the TJ Close lab at the University of
California, Riverside (Akhunov, Chin, Choi, Close, Fenton,
Kianian, Otto, Simons, Zhang). Phagemids were plated and
picked at the Clemson University Genomics Institute (CUGI)
(Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA
preparations, DNA sequencing and sequence analysis were
performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates
, Rambo, Main). The sequence has been trimmed to remove
vector sequence and contains a minimum of 100 bases of
phred value 20 or above. For more details on library
preparation and sequence analysis see
http://www.genome.clemson.edu/projects/barley. To order
this clone see http://www.genome.clemson.edu/orders. Also
see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
Genetically and physically anchored EST resources for
barley genomics. Barley Genetics Newsletter 31:29-30.

BASE COUNT 120 a 186 c 161 g 98 t 1 others
ORIGIN
Query Match 8.4%; Score 209.8; DB 12; Length 566;
Best Local Similarity 70.3%; Pred. No. 1e-32;
Matches 296; Conservative 0; Mismatches 122; Indels 3; Gaps 1;

QY 1411 AGAGGCCACTCGACGCTGGGCGGCTGGGAGGTGTGGGCGAGGCGGCTGGGCGGCATCC 1470
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 AGAAGCCCTCACTGACGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTGC 61
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1471 TCCACGAGGACGACGCTGGGCGGCTGGGAGGTGTGGGCGAGGCGGCTGGGCGGCATCC 1530
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62 TCCACGAGGAGGCTGTGGGCGGCTGGGAGGTGTTCACCGGCTGGGCGGTGGCGTGC 121
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1531 GCCCTCGCCCGGGGCGCTGCTGCTCAACGCTCGGCGACACCTTCATGAGGCGCTCTCCACG 1590
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122 GGCCTCGGAGGCGGCGCTGCTGCTCAACGCTCGGCGACACCTTCCTCGGCGCTCAACG 181
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1591 CCAGGTACCGGAGCTGCTGCTGACCGGCGGCTGCTCAACGACGCGGCGCTGCGGCTGC 1650
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 182 GCGCGCAGCTGCTGCTGCTCAACGCGGCGGCTGCTCAACGCGGCGGCGGCGGCTGC 241
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1651 TGGCTCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1710
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 242 TCACCTTCTTCTCAACCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 301
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1711 AC---GACCCACCGGAGGTGTACCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1767
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 302 CCATTGATGGAGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 361
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1768 CGCAGCGGCGCTACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 1827
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 362 CGCAGAGCACTACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1828 A 1828
Db 422 A 422

RESULT 10
BE454826 571 bp mRNA linear EST 22-OCT-2001
LOCUS
DEFINITION
HVSMEH0095D08f Hordeum vulgare 5-45 DAP spike EST library
HVCNDA0009 (5 to 45 DAP) Hordeum vulgare cDNA clone HVSMEH0095D08f,
mRNA sequence.
ACCESSION
BE454826
VERSION
BE454826.3 GI:16322121
KEYWORDS
EST.
SOURCE
Hordeum vulgare.
ORGANISM
Hordeum vulgare.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.
1 (bases 1 to 571)
Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Begum, D., Frisch, D., Yu
, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi, D.W., Fenton
, R.D., Close, S.J., Oates, R. and Main, D.
Development of a genetically and physically anchored EST resource
for barley genomics: Morex 5-45 DAP spike cDNA library
Unpublished (2001)
On Jul 26, 2000 this sequence version replaced gi:13189555.
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hg bases = 518
Seq primer: AATTAACCTCACTAAAGGG
High quality sequence stop: 556.

```

FEATURES  
source

Location/Qualifiers  
1. .571  
/organism="Hordeum vulgare"  
/cultivar="Morex"  
/db\_xref="taxon:4513"  
/clone="HVSMEH0095D081"  
/clone\_lib="Hordeum vulgare 5-45 DAP spike EST library  
HVCNDA0009 (5 to 45 DAP)"  
/tissue\_type="5-45 DAP Spike"  
/lab\_host="SOLR"  
/note="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI;  
Plants were grown in the greenhouse at the University of  
California, Riverside (Fenton, SJ Close, TJ Close). Whole  
spikes with awns trimmed were collected at 5, 10, 15, 20,  
30 and 45 DAP (Fenton). Total RNA was prepared from each  
pool, equal quantities of all six RNA pools were combined,  
poly(A) RNA was purified from the mixture, one primary  
unamplified cDNA library was made, and 1 million pfu were  
in vivo excised to give pBluescript SK(-) cDNA phagemids  
(Choi) in the TJ Close lab at the University of California,  
Riverside. Phagemids were plated and picked at the Clemson  
University Genomics Institute (CUGI) (Begum, Palmer,  
Frisch, Atkins and Wing). Plasmid DNA preparations, DNA  
sequencing and sequence analysis were performed at CUGI  
(Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main).  
The sequence has been trimmed to remove vector sequence  
and contains a minimum of 100 bases of phred value 20 or  
above. For more details on library preparation and  
sequence analysis see  
http://www.genome.clemson.edu/projects/barley. To order  
this clone see http://www.genome.clemson.edu/orders  
see close TJ, Wing R, Kleinbofs A, Wise R (2001)  
Genetically and physically anchored EST resources for  
barley genomics. Barley Genetics Newsletter 31:29-30.  
(http://wheat.pw.usda.gov/gbgpages/bgn/31/cover.html)"

BASE COUNT 92 a 202 c 190 g 87 t  
ORIGIN  
Query Match 8.1%; Score 203.8; DB 10; Length 571;  
Best Local Similarity 69.3%; Pred. No. 1.7e-31;  
Matches 293; Conservative 0; Mismatches 127; Indels 3; Gaps 1;  
QY 775 TGGTGTTCAGCGCGGCTGTGACGGGAGAGATCCCGTCGAGTTCATATGCG 834  
Db 114 TCGTGTTCAGCGCGGCTGTGACGGGAGAGATCCCGTCGAGTTCATATGCG 173  
QY 835 CGGCGGAGAGACCCCGGCTGTGACGGGAGAGATCCCGTCGAGTTCATATGCG -- 892  
Db 174 CAGCGGAGAGACCCCGGCTGTGACGGGAGAGATCCCGTCGAGTTCATATGCG 233  
QY 893 -CGTGGGCGGCGGAGAGATCCCGTGTGACGGGAGAGATCCCGTCGAGTTCATATGCG 951  
Db 234 TCGCGGCTTCGCGGCGGAGAGATCCCGTGTGACGGGAGAGATCCCGTCGAGTTCATATGCG 293  
QY 952 GCGAGCGCTTCCTGCGGCTTCACCGGATCGAGCGGCGCTGTGAGGAGCGCG 1011  
Db 294 GCGAGCGCTTCCTGCGGCTTCACCGGATCGAGCGGCGCTGTGAGGAGCGCG 353  
QY 1012 ACCGCTGAGGAGCGCTTCCTGCGGCTTCACCGGATCGAGCGGCGCTGTGAGGAGCGCG 1071  
Db 354 ACCGCTGAGGAGCGCTTCCTGCGGCTTCACCGGATCGAGCGGCGCTGTGAGGAGCGCG 413  
QY 1072 GCGGCGGAGAGATCGCGCTACCGGATCGAGCGGCGCTGTGAGGAGCGCG 1131  
Db 414 GCGTGGGAGAGATCGCGCTACCGGATCGAGCGGCGCTGTGAGGAGCGCG 473  
QY 1132 CGTGGAGGAGAGATCGCTTCGCGGCTTCACCGGATCGAGCGGCGCTGTGAGGAGCGCG 1191  
Db 474 CGTGGAGGAGAGATCGCTTCGCGGCTTCACCGGATCGAGCGGCGCTGTGAGGAGCGCG 533  
QY 1192 GCG 1194  
Db 534 TCG 536

RESULT 11

BQ606563  
LOCUS  
DEFINITION BRY.2428 wheat EST endosperm library Triticum aestivum CDNA 5',  
mRNA sequence.  
ACCESSION BQ606563  
VERSION BQ606563.1 GI:21555890  
KEYWORDS EST.  
SOURCE bread wheat.  
ORGANISM Triticum aestivum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae  
; Triticeae; Triticum.  
REFERENCE 1 (bases 1 to 567)  
AUTHORS Clarke,B., Lambrecht,M. and Rhee,S.  
TITLE Assessing the utility of Arabidopsis genomic information for  
interpreting wheat EST sequences  
JOURNAL Unpublished (2002)  
COMMENT Contact: Lambrecht M  
The Arabidopsis Information Resource  
Carnegie Institution of Washington, Dept. of Plant Biology  
260 Panama Street, Stanford, CA 94305, USA  
Tel: 1 650 325 1521 x 251  
Fax: 1 650 325 3748  
Email: rhees@cuma.stanford.edu.  
FEATURES  
source  
1. .567  
/organism="Triticum aestivum"  
/cultivar="Wyuna"  
/db\_xref="taxon:4565"  
/clone\_lib="wheat EST endosperm library"  
/tissue\_type="endosperm"  
/dev\_stage="developing endosperm tissue 8, 10 and 12 DPA  
(days post anthesis)"  
/note="Vector: Bluescript II SK(-)"

BASE COUNT 80 a 200 c 203 g 83 t 1 others  
ORIGIN  
Query Match 8.0%; Score 200; DB 14; Length 567;  
Best Local Similarity 66.7%; Pred. No. 1e-30;  
Matches 354; Conservative 0; Mismatches 156; Indels 21; Gaps 4;  
QY 775 TGGTGTTCAGCGCGGCTGTGACGGGAGAGATCCCGTCGAGTTCATATGCG 834  
Db 46 TCGTGTTCAGCGCGGCTGTGACGGGAGAGATCCCGTCGAGTTCATATGCG 105  
QY 835 CGGCGGAGAGACCCCGGCTGTGACGGGAGAGATCCCGTCGAGTTCATATGCG 894  
Db 106 CAGCGGAGAGACCCCGGCTGTGACGGGAGAGATCCCGTCGAGTTCATATGCG 165  
QY 895 TGG-----GGCGGGGCGGAGAGATCCCGTCGAGTTCATATGCG 942  
Db 166 TCGCGGCTTCCTGCGGCGGAGAGATCCCGTCGAGTTCATATGCG 225  
QY 943 CGTGGAGAGAGCGCTTCCTGCGGCTTCACCGGATCGAGCGGCGCTGTGAGGAGCGCG 1002  
Db 226 CGTGGAGAGAGCGCTTCCTGCGGCTTCACCGGATCGAGCGGCGCTGTGAGGAGCGCG 285  
QY 1003 AGAGAGCGGCGGCTTCCTGCGGCTTCACCGGATCGAGCGGCGCTGTGAGGAGCGCG 1062  
Db 286 CCAAGCGGCTTCCTGCGGCTTCACCGGATCGAGCGGCGCTGTGAGGAGCGCG 345  
QY 1063 CGAGAGCGGCGGCGGAGAGATCCCGTCGAGTTCACCGGAGAGATCCCGTCGAGTTCGCGGT 1122  
Db 346 CGAGAGCGGCGGCTTCCTGCGGCTTCACCGGATCGAGCGGCGCTGTGAGGAGCGCG 405  
QY 1123 CCAAGCTCCCGTGGAGAGAGATCCCGTCGAGTTCACCGGATCGAGCGGCGCTGTGAGGAGCGCG 1182  
Db 406 GCAAGCTCCCGTGGAGAGAGATCCCGTCGAGTTCACCGGATCGAGCGGCGCTGTGAGGAGCGCG 465  
QY 1183 CGAGAGAGAGCGGCTTCCTGCGGCTTCACCGGATCGAGCGGCGCTGTGAGGAGCGCG 1242

```

Db 466 GCAG-----GTGTCGACTTCTCGTCCGCTCTCGGCGAGGAGTAC---CGTGACA 516
QY 1243 TGGCGAGGTGTACTTCGCGTACTGCGACGAGATGAGCGCGCTGTGCGCTGG 1293
Db 517 TGGGAGAGGTGTGGCANGAGTACTGCGACGAGATGACGCGCTGGCGCTGG 567

RESULT 12
BG350197 655 bp mRNA linear EST 01-MAR-2001
LOCUS 084C04 Mature tuber lambda ZAP Solanum tuberosum cDNA, mRNA
DEFINITION
ACCESSION BG350197
VERSION BG350197.1 GI:13178939
KEYWORDS EST.
SOURCE Solanum tuberosum
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1. (bases 1 to 655)
NIELSEN,K.L., CROOKSHANKS,M., EMMERSEN,J. and WELINDER,K.G.
EST-sequencing of mature potato tuber (Var. Kuras)
JOURNAL Unpublished (2000)
COMMENT Contact: Karen G. Welinder
Institut for bioteknologi
Aalborg Universitet
Sohngaardsholmsvej 49, 9000 Aalborg, Denmark
Tel: +45 96358467
Fax: +45 98141808
Email: kgw@bio.auc.dk
Sequenced from the 5' end.
High quality sequence stop: 655
POLYA-No.

FEATURES
    source
    1. 655
        /organism="Solanum tuberosum"
        /cultivar="Field grown Kuras"
        /db_xref="taxon:4113"
        /clone_lib="Mature tuber lambda ZAP"
        /tissue_type="tuber"
        /notes="Vector: Lambda ZAP"
BASE COUNT 224 a 129 c 120 g 180 t 2 others
ORIGIN
    Query Match 7.9%; Score 196.8; DB 12; Length 655;
    Best Local Similarity 63.3%; Pred. No. 4.5e-30;
    Matches 300; Conservative 0; Mismatches 174; Indels 0; Gaps 0;

QY 1360 TCTTCAGCGCAAGACTCCATCATCGGCTCACTACTACCGCGTGCAGAGCGCAC 1419
Db 9 TATTCGAAGAAATGATTTCGATAATGAGACTCACTACTATCCACCATGTCAGAAACGG 68
QY 1420 TCACACGCTGGGACCGGTCGCGACTGCGACCCACCTCTCCATCCATCTCCACG 1479
Db 69 AGCTCACTTAGGAACGGCGCTCATGTGTGACCCCAACATTAACAATTCCTCAAG 128
QY 1480 ACCACCTGGCGGCTGGAGGTGTGGCGAGGCGGTGGCGGCATCGCGCTCGCC 1539
Db 129 ATTGTGTGGCGGACTTCAAGTTTGTGGATGATGAATGCGGTTCATCGTCAATTT 188
QY 1540 CCGGGGCGCTCGTCTCAACGTGGGACACCTTTCATGGCGCTCTCAACGCCAGTAC 1599
Db 189 TCAATGCAATTTGTGTTAAATAGGCGACACATNTATGGCGCTATCAATAGGAGATCA 248
QY 1600 GCAGTGTCTGCACCGGCGGTGTCAACAGACGCGCGCTCGCGCTCGCTGGCGCTCT 1659
Db 249 AAAGTTGCGTCACAGACGATAGTAACAACAGACTCTCTAGAAAATCACTTGCTTCT 308
QY 1660 TCTCTGCGCGGAGATGACACGCTGTGGCGCGCGCGGAGAGCTGGTGCAGCACCC 1719
Db 309 TTCTTTGTCGAACAAAGATNAGGTGGTGAGCCCAACAAATGAATTGGTGAATCCAA 368

```

```

QY 1720 ACCGAGGTGTACCGGACTTCACGTGCGGGCGCTGCTGGACTTCACGACGCCACT 1779
Db 369 ACCCTCGAATATATCTGATTTCACATGCGCTACTCTCTTGAATTTACTCAAAACATT 428
QY 1780 ACAGGCGGACATCGCAGCGTCTCAGGCTTCTCGACTGCTTAATCATCATC 1833
Db 429 ATAGAGCTGATATGACACTCTTCAAAACATTCTCAAATNGGCTCATGATCAAC 482

RESULT 13
BE422495 517 bp mRNA linear EST 24-JUL-2000
LOCUS WHE0056_E02_I04Zs Wheat endosperm cDNA library Triticum aestivum
DEFINITION
ACCESSION BE422495
VERSION BE422495
KEYWORDS EST.
SOURCE BE422495.1 GI:9420410
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Triticum.
1. (bases 1 to 517)
ALTENBACH,S., ANDERSON,O.D., CHAO,S., GALILI,G., HAN,P.S., HSIA
,C.C., KANG,Y., LAZO,G.R., MILLER,R., RAUSCH,C.J., SEATON,C.L. and
TONG,J.C.
The structure and function of the expressed portion of the wheat
genomes - Endosperm cDNA library
JOURNAL Unpublished (2000)
COMMENT Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: Stratagene SK primer.

FEATURES
    source
    1. 517
        /organism="Triticum aestivum"
        /cultivar="Cheyenne"
        /db_xref="taxon:4565"
        /clone="WHE0056_E02_I04"
        /clone_lib="Wheat endosperm cDNA library"
        /tissue_type="Endosperm"
        /dev_stage="5 to 30 days post anthesis seed"
        /lab_host="E. coli SOLR"
        /notes="Vector: Lambda ZAP II, excised phagemid; Site_1:
        EcoRI; Seeds collected, endosperm isolated, and RNA
        prepared by Susan Altenbach. Library constructed by
        Stratagene, Inc. Plasmid DNA preparations and DNA
        sequencing were performed in the OD Anderson lab."
BASE COUNT 96 a 149 c 161 g 111 t
ORIGIN
    Query Match 7.6%; Score 191.2; DB 10; Length 517;
    Best Local Similarity 71.5%; Pred. No. 6.6e-29;
    Matches 266; Conservative 0; Mismatches 103; Indels 3; Gaps 1;

QY 1460 GCTCACCATCTCCACGAGCACCGTCCGCGCGCTGGAGGTGTGGCGAGGCGGTG 1519
Db 517 GCTGACGCTGCTCCACGAGCACGCTCGCGGGCTGGAGGTGTTCACCGGCGCGGTG 458
QY 1520 GCGGCGCATTCGCGCTCGCCCGGGCGCTCGTCTGCTCAACGTGGGACACCTTCATG 1579
Db 457 GCAGCGGCTGCGCGCGCGGAGCGACCGCTTCTGCTCAACATCGGCGACACTTCTCG 398
QY 1580 GCTCTCCACGCGCAGGTACCGGAGCTCGCTGACCGGCGGTGCTCAACGACGCGCC 1639
Db 397 GCTCACCAACGCGCGCCACATCATGCTCCACCGCGCGTGTCTCAACGAGCGCTGCG 338

```

```

QY 1640 TGGCGCTCGTGGCTCTTCTCTGCCCCGAGATGACAGTGGTGGCGCCGCCGGA 1699
Db 337 CGCGAGTGGCTGACCTCTCTCTCAACCCGAGCTGGACCGCTGTTGGCGCCGCCGCG 278
QY 1700 GGAGCTG---GTGACAGACACACACCCCGAGGCTGTACCCGAGCTTACGTTGGCGGCGCT 1756
Db 277 CGAGCTGCTCGCGCTGCGACGCGCCGCCGCGCTGTACCCGAGCTTACGTTGGCTGAAT 218
QY 1757 GTGGAGCTTCAGAGCGCCGCTACAGGCGCCGACATGCGGACGCTTACAGCGCTTTCGGA 1816
Db 217 CTTCGAGTTCAGAGAGGAGTACCGCTCTGACTGGAGAGCCCTGGATGCATGCGCTC 158
QY 1817 CTGCGTAAATCA 1828
Db 157 GTGATCAATCA 146

RESULT 14
BE414214
LOCUS BE414214 686 bp mRNA linear EST 24-JUL-2000
DEFINITION SCU007.E01.R990714 ITFC SCU Wheat Endosperm Library triticum
aestivum cDNA clone SCU007.E01, mRNA sequence.
ACCESSION BE414214
VERSION BE414214.1 GI:9412060
KEYWORDS EST.
SOURCE bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 686)
AUTHORS Anderson,O.A., Appels,R., Bailey,P., Blake,T., Close,T., Cloutier
,S., Dubcovsky,J., Feuillet,C., Gale,M., Graner,A., Gustafson,P.,
Herrmann,R.G., Holton,T., Jacquemin,J.M., Jia,J., Joudrier,P.,
Langridge,P., Lazo,G.R., Lin,J.J., McGuire,P., Ogihara,Y.,
Pecchioni,N., Qualset,C., Schuch,W., Selvaraj,G., Shariflou,M.,
Sorrells,M., Warburton,M. and Wenzel,G.
TITLE International Triticale EST Cooperative (ITEC): Production of
Expressed Sequence Tags for Species of the Triticeae
JOURNAL Unpublished (2000)
COMMENT Contact: Holton T
Centre for Plant Conservation Genetics, Southern Cross University
PO Box 157, Lismore NSW 2480 AUSTRALIA
Tel: 61 2 6620 3409
Fax: 61 2 6622 2080
Email: tholtonscu.edu.au
International Triticale EST Cooperative (ITEC)
http://wheat.pw.usda.gov/genome.
FEATURES
source
1. 686
/organism="Triticum aestivum"
/cultivar="Wyuna"
/db_xref="taxon:4565"
/clone="SCU007.E01"
/tissue_type="endosperm"
/notes="Vector: Bluescript II SK(-)"
BASE COUNT 112 a 232 c 232 g 105 t 5 others
ORIGIN
Query Match 7.5%; Score 187; DB 10; Length 686;
Best Local Similarity 68.1%; Pred. No. 4.6e-28;
Matches 294; Conservative 0; Mismatches 126; Indels 12; Gaps 2;

QY 775 TGTGTTTCAGCGCGGCTGTCAGCGGGCAGACGAGATCCCTCCGATTCATAGGC 834
Db 159 TCGTGTTCAGCGCGGCTGTCAGCGGGCAGACGAGATCCCTCCGATTCATAGGC 218
QY 835 CGCGGAGAGAGAGCCCGCTCGCTGGAGGAGTGGAGAGCTGGAGGAGTGGAGGAG 894
Db 219 CAGCCAGAGAGAGCCCGCTCGCTGGAGGAGTGGAGGAGTGGAGGAGTGGAGGAG 278

```

```

QY 895 TGG-----GGCGGGGCGGAGAGTGTGTCGTGG--TCCGCGAGTGGGGAGG 942
Db 279 TCGCGGGTTCCTTGGCGGGGAGCGGCTGCCACTTGGCGGCTCCGTGACTGGCGCGG 338
QY 943 CFTGGCAGAGCAGCGGCTTCTTCTGCTGTTTAAACCGCATCGAGGCGGCTGCTGG 1002
Db 339 CFTGGCAGAGCAGCGGCTTCTTCCAGTGTGACACCGAGGCTGGACCGGCGCTGCG 398
QY 1003 AGAGGCGCAGCGGTGTCATGAGCGCTTCTTACGCTCCGCTGGGGGAGAGCAGCGGG 1062
Db 399 CCAAGCGCTTACCGTGTGCGAGCGCTTCTTACGCTCCGCTCGCGGAGAGCAGCGG 458
QY 1063 CGCAGGCGCGCGGGGAGAGCTGCGGCTAGCCAGAGCTTACGCGGCGCTTGGCT 1122
Db 459 CGCAGGCGCGCGCTGCGGAGACACCGGTAGCGGCGGCTGCTGGGGCGGGTCGGA 518
QY 1123 CCAAGCTCGCTGGAGGAGACGCTGCTTCCGCTACTCATCTGGAGATGAGAGG 1182
Db 519 GCAAGTCCCTGGAGGAGACCGCTGCTTCACTGTCTCGCGCGCGCGGANGGCC 578
QY 1183 GCGAGGAGGCGG 1194
Db 579 GCAAGTCTGCTG 590

RESULT 15
BQ402912
LOCUS BQ402912 700 bp mRNA linear EST 22-MAY-2002
DEFINITION GA_Ed0052C12f Gossypium arboreum 7-10 dpa fiber library Gossypium
arboreum cDNA clone GA_Ed0052C12f, mRNA sequence.
ACCESSION BQ402912
VERSION BQ402912.1 GI:21090599
KEYWORDS EST.
SOURCE Gossypium arboreum.
ORGANISM Gossypium arboreum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
REFERENCE 1 (bases 1 to 700)
AUTHORS Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
TITLE An integrated analysis of the genetics, development, and evolution
of the cotton fiber
JOURNAL Unpublished (2000)
COMMENT Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 854 656 7288
Fax: 854 656 4293
Email: rwing@clemson.edu
Total High Quality bases = 572
Seq primer: TAATCACTCACTATAGG
High quality sequence start: 7
High quality sequence stop: 680.
Location/Qualifiers
1. 700
/organism="Gossypium arboreum"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ed0052C12f"
/tissue_type="Fibers isolated from bolls harvested 7-10
dpa"
/lab_host="E. coli"
/notes="Vector: pBR-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 194 a 148 c 171 g 186 t 1 others
ORIGIN
Query Match 7.4%; Score 185.6; DB 14; Length 700;
Best Local Similarity 59.9%; Pred. No. 8.8e-28;
Matches 371; Conservative 0; Mismatches 230; Indels 18; Gaps 3;

```

Search completed: March 26, 2003, 07:52:50  
Job time : 2570.09 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2003, 04:38:30 ; Search time 61.8613 Seconds  
(without alignments)  
12398.686 Million cell updates/sec

Title: US-09-702-134-7212\_COPY\_28000\_30500

Perfect score: 2501  
Sequence: 1 atgtcaagataaaattggc.....ctatcaatggtgtcttcaat 2501

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents\_NA:\*

- 1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	292.4	11.7	1259	2	US-08-553-367A-3
2	292.4	11.7	1259	4	US-09-295-306-3
3	292.4	11.7	1259	4	US-09-734-719-3
4	271	10.8	1490	2	US-08-553-367A-5
5	271	10.8	1490	4	US-09-295-306-5
6	271	10.8	1490	4	US-09-734-719-5
7	226.4	9.1	1395	2	US-08-553-367A-1
8	226.4	9.1	1395	4	US-09-295-306-1
9	226.4	9.1	1395	4	US-09-734-719-1
10	94.4	3.8	1926	4	US-09-249-585A-2
11	94.4	3.8	2580	3	US-09-050-863-2
12	94.4	3.8	2580	4	US-09-359-081-2
13	94.4	3.8	5452	2	US-09-130-114-1
14	94.4	3.8	9600	4	US-08-910-647-1
15	94.4	3.8	9600	4	US-09-620-925-1
16	94.4	3.8	10596	1	US-07-884-811-15
17	94.4	3.8	10596	1	US-07-885-971-15
18	94.4	3.8	10596	1	US-08-087-783A-15
19	94.4	3.8	10596	1	US-08-194-088B-15
20	94.4	3.8	10596	2	US-08-194-087-15
21	94.4	3.8	10596	5	PCT-US93-04648-15
22	94.2	3.8	1181	2	US-08-632-598-2
23	94.2	3.8	1181	4	US-09-231-240-2
24	90	3.6	1926	4	US-09-249-585A-4
25	90	3.6	1931	2	US-09-130-114-2
26	79.8	3.2	1505	1	US-07-915-246-1
27	78.2	3.1	2793	1	US-08-209-747-1

28	78.2	3.1	2793	1	US-08-458-298-1	Sequence 1, Appli
29	77.2	3.1	672	4	US-09-454-034-9	Sequence 9, Appli
30	73.8	3.0	174	2	US-08-553-367A-12	Sequence 12, Appli
31	73.8	3.0	174	4	US-09-295-306-12	Sequence 12, Appli
32	73.8	3.0	174	4	US-09-734-719-12	Sequence 12, Appli
33	72.6	2.9	888	4	US-08-765-907A-2	Sequence 2, Appli
34	72.6	2.9	2888	4	US-08-765-907A-1	Sequence 1, Appli
35	64.6	2.6	2338	1	US-08-425-069-1	Sequence 1, Appli
36	64.6	2.6	2338	1	US-08-317-844B-1	Sequence 1, Appli
37	64.2	2.6	7218	1	US-08-232-463-14	Sequence 14, Appli
38	64	2.6	1228	2	US-08-823-986A-1	Sequence 1, Appli
39	64	2.6	1228	3	US-08-784-385-1	Sequence 1, Appli
40	64	2.6	1270	5	PCT-US95-10403-1	Sequence 1, Appli
41	64	2.6	1663	2	US-08-823-986A-3	Sequence 3, Appli
42	64	2.6	1663	3	US-08-784-385-3	Sequence 3, Appli
43	64	2.6	1703	5	PCT-US95-10403-3	Sequence 3, Appli
44	62.6	2.5	470	4	US-09-615-192A-149	Sequence 149, App
45	62.6	2.5	2830	2	US-09-010-928B-1	Sequence 1, Appli

## ALIGNMENTS

## RESULT 1

US-08-553-367A-3  
; Sequence 3, Application US/08553367A  
; Patent No. 5939539  
; GENERAL INFORMATION:  
; APPLICANT: Theodor LANGE et al.  
; TITLE OF INVENTION: GA 20-OXIDASE GENE SEQUENCES (As Amended)  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.  
; STREET: 2033 K Street, N.W., Suite 800  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/553,367A  
; FILING DATE: No. 5939539 September 27, 1995  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lee Cheng  
; REGISTRATION NUMBER: 40,949  
; REFERENCE/DOCKET NUMBER: 49/FDA.5M2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-721-8200  
; TELEFAX: 202-721-8250  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1259 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA to mRNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Arabidopsis thaliana  
; IMMEDIATE SOURCE:  
; CLONE: pAC2301  
US-08-553-367A-3

Query Match	11.7%;	Score 292.4;	DB 2;	Length 1259;
Best local similarity	60.0%;	Pred. No. 2.2e-54;		
Matches 532;	Conservative	0;	Mismatches 336;	Indels 18; Gaps 2;
QY 939	GAGGCGTCGAGAGGACACGGCTTCTTCTGCTGTTAAACCGGATCGAGCGGCGCTG	998		
Db 264	GAGGCGTGAAGAAGCACGGTTTCTCTCGTGGTCAATCAGCGATCAGCGAGAGCTT	323		
QY 999	CTGGAGAGGCGCACGGGTGCATGAGCGCTTCTTCCGCTCCGCTGGGGGGAAGAAGC	1058		
Db 324	ATATCAGACGCTCATGAATACACGAGCGCTCTTTTGATATGCCCTCTCCGAAACAG	383		
QY 1059	CGGGCGCAGCGCGCGCGGGGAGAGCTGCGGCTACGCCAGCAGCTTCACGGGCGCTTC	1118		
Db 384	AGGGTCTTTAGAAAAATCCGGTGAGAGTGTGGCTTAAGCAAGCAGTTTACCGGACGCTT	443		
QY 1119	GGTCCAAAGCTCGCGTGGGAAGAGAGCGTGCCTCCGCTACTCATCGCTGGAGATGA	1178		
Db 444	TCACCAAGCTTCATGGGAAGAGACCTTCTTTTCGGT-----TTTGGCGAGC	494		
QY 1179	GAGGCGAGGAGGCGCTGGTGAGTACCTGGTGGGAAGCTCGGGGCGAGCAGCGGCG	1238		
Db 495	ATGAGCGCTCAAAATCCGTTCAAGATTACTTTCGATCGCTTGGGACATGGGTTTCAG	554		
QY 1239	CGGCTGGCGAGGTGTACTCGCGCTACTGCGACAGATGAGCGGCGCTCTCCCTGGAGCTG	1298		
Db 555	CCATTTGGGAAGGTGTATCAAGAGTATTGTGAAGCAATGAGTTCTCTATCATTTAAGATC	614		
QY 1299	ATGAGGTGTCTCGGGGAGAGCTGGGACATCGTCGGAGACGGCGGCACACTCTCCGCGCA	1358		
Db 615	ATGGAGCTTCTCGGGCTAAGTTTAGC-----GTAAACGGGCTACTTTAGAGAG	665		
QY 1359	TTCTTCCAGCGAACGACTCCATCATGGCGCTCAACTACTTACCGGCGTGCAGAGGCCA	1418		
Db 666	TTTTTCGAAGAAGAACGATTCAATATGAGACTGAATTACTTACCTCCATGTATAAACCA	725		
QY 1419	CTCGACACGCTGGCGACCGGTGCGCACTGCGACCCCACTCGCTCAACATCCTCCACAG	1478		
Db 726	GATCTCACACTAGGAACAGAGACTCATTTGATCCACATCTCTTACCATCTTCAACAA	785		
QY 1479	GACCACTGCGCGGCTCGAGGTGTGGCGAGGGGGTGGCGGCATCGGCCCTCGC	1538		
Db 786	GACCATGTAAATGGCCCTCAAGTCTTTGGAATAATCAATGGCGCTCCATTCGTCCTCAAC	845		
QY 1539	CCCGGGCGCTCGCTCGTCAACCTGCGGCACACTTTCATGGCGCTTCCAAAGCCACAGGTAC	1598		
Db 846	CCCAAGGCGCTTTGTGTCAATATCGCGCATACTTTCATGGCTCTATCGACGATAGTATC	905		
QY 1599	CGCACTGCTCGACCGGGGGTGTCTCAACAGACGGCGCTCGCGGCTCGCTGGCGCTTC	1658		
Db 906	AAGAGCTCTTTCACCGCGGCTGTGAACAGAGCGGAGGAAGTCTCTTTCATTC	965		
QY 1659	TTCTCTCCCGGAGATGACACGGTGTGTGGCGCGCGGAGAGCTGGTCGAGACCAAC	1718		
Db 966	TTCTTGTGTGCGAAATAAGACAGATGTAGTGAGCCACCGAGAGAGCTTTTGGACAGATC	1025		
QY 1719	CACCGAGGGTGTACCCGAGTTCACGTGGCGGGCGCTGTGGACTTCACGACGCGCAC	1778		
Db 1026	ACATCAAGAAGATACCTGACTTTCACATGGTCTATGTTCCTTGGTTCACTCAGAAACAT	1085		
QY 1779	TACAGGGCGGACATGCGCAGGTTCAGGCGCTTCTCCGACTGGCTTA	1824		
Db 1086	TATAGACGACATGACATCTCAAGCCTTTTCAGATTGGCTCA	1131		

## RESULT 2

```

RESOLUTION 2
US-95-295-306-3
; Sequence 3, Application US/09295306
; Patent No. 6198021
; GENERAL INFORMATION:
; APPLICANT: Theodor LANGE et al.
; TITLE OF INVENTION: GA 20-OXIDASE GENE SEQUENCES (As Amended)
; NUMBER OF SEQUENCES: 19

```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
STREET: 2033 K Street, N.W., Suite 800
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/295,306
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/553,367
FILING DATE: No. 6198021ember 27, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Lee Cheng
REGISTRATION NUMBER: 40,949
REFERENCE/DOCKET NUMBER: 49/DIV-FD4.5M2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-721-8200
TELEFAX: 202-721-8250
TELEX:
INFORMATION FOR SEO ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1259 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA Lo mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Arabidopsis thaliana
IMMEDIATE SOURCE:
CLONE: PAT2301
US-09-295-306-3

Query Match 11.7%; Score 292.4; DB 4; Length 1259;
Best local Similarity 60.0%; Pred. No. 2.2e-54;
Matches 532; Conservative 0; Mismatches 336; Indels 18; Gaps

QY 939 GAGCGCTGGAGAGCAGCGCTCTCTCTCTGTGTGTAAACACAGGCATCAGAGCGCGCTG 998
|||||
Db 264 GAGCGCTGTAAAGCAGCGTTCTCTCTCTGTGTCAATCAGGCATCAGCAGAGGCTT 323
|||||
QY 999 CTGGAGAGCGCCACCGGTGCATGAGCAGCCTTCTTTCACGCTCCGCTGGGGGAGAGCAG 1058
|||||
Db 324 ATATCAGACGCTCATGAATACAGCGCGCTCTTTGATACGCTCTCTCCGAAAAACAG 383
|||||
QY 1059 CGGGCGCAGCGCGCGGGGAGAGCTGCGGCTAGCGCAGAGCTTACGCGGGCGCTTC 1118
|||||
Db 384 AGGTTCTTAGAAATCCGCTGAGAGTGTGGCTACGAGAGAGTTTACCAGCGCTTC 443
|||||
QY 1119 GCGTCAAGCTCGCTGGAGAGAGCGCTGTCTTCGGTATCATCGGCTGGAGATGAA 1178
|||||
Db 444 TCCACCAAGCTTCCATGGAAGAGAGACCCCTTTCTTTCGGT-----TTTCCGACGAC 494
|||||
QY 1179 GAGGGCAGAGGGCGTGGTGAGTACTGTGTGGGNAAGCTCGGGGGAGACAGGGGG 1238
|||||
Db 495 ATGAGCGCGTCAAAATCCGTTAAGATTACTTCTTGCGATGCGTTGGGACATGGGTTT 554
|||||
QY 1239 CGGCTGGCGAGGTACTCTCGCGCTACTGCCACAGATGAGCGCGCTGTGCTGGAGCTG 1298
|||||
Db 555 CCAATTGGGAGGTATCAAGAGTATGTGAACATGAGTTCTCTATCATTTAGATGC 614
|||||
QY 1299 ATGAGGTGCTCGGGAGAGCGCTGGGCATCTGTGGAGACCGGCCCACTACTTCCGGGA 1358
|||||
Db 615 ATGAGCTTCTCGGCTAAGTTTGGC-----GTAAACGGGACTACTTTAGAGAG 665
|||||

```

QY	1359	TTCTCTCAGCGCAACGACTCCATCATCTGCGGCTCAACTACTACCGGGGTCCGAGAGCCA	1418
Db	666	TTTTTCGAAGAAAGGATTCAATATATGAGACTGAAATATACCTCCATGATATAAACCA	725
QY	1419	CTCGACACGCTGGGACACGGTCCGCGACTGGACCCCAACCTCGCTCAACATCTCCACCAG	1478
Db	726	GATCTCACACTAGGACAGACACCTCAATGTGATCCAACTCTTTACATCTCTTCACTTCA	785
QY	1479	GACCACGTCGGCGCTCTGGAGTGTGGCGGAGGGGGTGGCGGCATCGCCCTGCG	1538
Db	786	GAACATGTAAATGGCTTCAAGTCTTTTGGAAATCAATGCGCTCCNTGTCCTCCAC	845
QY	1539	CCCGGGCGCTCGTGGTCAACGTCGCGACACCTTCATGCGCTCTCCACGCCAGGTAC	1598
Db	846	CCCAAGGCTTTTGGTCAATACGGCGATCTTTTCATGGCTCTATCGAACGTAGATAC	905
QY	1599	CGCAGCTGCTCCACCGGGGGTCTGTCACAGCAAGCGGCTTCGCGCTCGTGCCTTC	1658
Db	906	AGAGCTGTCTTCACCGGGGGTGTGTGACAGCAGAGGAGAGAASTCACTTCGATTC	965
QY	1659	TTCTCTTGCOCGGAGATGACACGTTGGTGGCCGCCCGGAGAGCTGTGCACGACAC	1718
Db	966	TTCTTTGTGTCCGAAAAAGACAGAGTAGTGACGCCACGACGAGAGCTTTTGACAGCATC	1025
QY	1719	CACCCGAGGGTGTACCCGGACTTCACGTGGCGGGGCGCTCGTGGACTTCAGCAGCGCCAC	1778
Db	1026	ACATCAGAGAGATACCTTGACTTCATCAGCTGTATGTTCTCTTGAGTTCACTCAGAAACAT	1085
QY	1779	TACAGGGCGCAATGCGCAGCTTCAGGCTTCTCCGACTGCTTAA	1824
Db	1086	TATAGACAGACATCAACATCTTCCAAAGCTTTTTCAGATTGCTCA	1131

### RESULT 3

```

US-09-734-719-3
; Sequence 3, Application US/09734719
; Patent No. 6455675
; GENERAL INFORMATION:
; APPLICANT: Theodor LANGE et al.
; TITLE OF INVENTION: GA 20-OXIDASE GENE SEQUENCES (As Amended)
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Wenderoth, Lind & Ponack, L.L.P.
; STREET: 2033 K Street, N.W., Suite 800
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09734,719
; FILING DATE: 13-Dec-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/295,306
; FILING DATE: April 21, 1999
; APPLICATION NUMBER: 08/553,367
; FILING DATE: No. 6455675, March 27, 1995
; APPLICATION NUMBER: PCT/EP94/01664
; FILING DATE: May 24, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee Cheng
; REGISTRATION NUMBER: 40,949
; REFERENCE/DOCKET NUMBER: 2000-1678/LC/01784
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-721-8200
; TELEFAX: 202-721-8250
; TELEX: <Unknown>

```

```

Db 1026 ACATCAAGAGATACCCCTGATTCACATGCTATGTTCTGATTCCTACTCAGAAACAT 1085
Qy 1779 TACAGGCGGACGACGCGCTTACAGGCTTCTCCGACTGGCTTA 1824
Db 1086 TATAGAGCAGACATGAACACTCTCCAGCCCTTTTCAGATTGGCTCA 1131

RESULT 4
US-08-553-367A-5
; Sequence 5, Application US/08553367A
; Patent No. 5939539
; GENERAL INFORMATION:
; APPLICANT: Theodor LANGE et al.
; TITLE OF INVENTION: GA 20-OXIDASE GENE SEQUENCES (As Amended)
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
; STREET: 2033 K Street, N.W., Suite 800
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/553.367A
; FILING DATE: No. 5939539 September 27, 1995
; CLASSIFICATION: 800
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee Cheng
; REGISTRATION NUMBER: 40,949
; REFERENCE/DOCKET NUMBER: 49/FD4.5M2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-721-8200
; TELEFAX: 202-721-8250
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1490 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHEtical: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Arabidopsis thaliana
; IMMEDIATE SOURCE:
; CLONE: PAT2353
; US-08-553-367A-5

Query Match 10.8%; Score 271; DB 2; Length 1490;
Best Local Similarity 56.5%; Pred. No. 1e-49;
Matches 593; Conservative 0; Mismatches 435; Indels 21; Gaps 4;

Qy 778 TGTTCAGCGCGGCTGCTGAGCGGACGAGATCCCGTCGCACTTCATATGGCCGG 837
Db 175 TCCTTAACCTCTCTCTTAACCTCCAAATCCCAAAACCAATTCATTGGCCAG 234
Qy 838 CGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 897
Db 235 ACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 291
Qy 898 GGGCGGGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 957
Db 292 CAAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 351

```

```

Qy 958 GCTTCTCTCTGTTGTTAAACACGATCAGAGCGCGCTCTGAGAGAGCGCACCGGT 1017
Db 352 GCTTCTCTCTGTTGTTAAACACGATCAGAGCGCGCTCTGAGAGAGCGCACCGGT 411
Qy 1018 GCATGAGAGCGCTCTCTTCAACGCTCGCGTGGGGAGAGAGAGAGAGAGAGAGAGAG 1077
Db 412 TGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 471
Qy 1078 GGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1137
Db 472 GTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 531
Qy 1138 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1197
Db 532 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 585
Qy 1198 GTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1257
Db 586 AAGATTACTTTTCCGATACATTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 642
Qy 1258 CGCGCTACTGCGACGAGATGAGCGCGCTCTGCTGAGAGAGAGAGAGAGAGAGAGAG 1317
Db 643 AAGACTATTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 702
Qy 1318 GCGTGGGATCGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1377
Db 703 GTTAGGCGGTAACCGAG-----ACTATTCCGAGAGAGAGAGAGAGAGAGAGAGAG 753
Qy 1378 CATCATGCGGCTCACTACTACTACCGCGGTCGAGAGAGAGAGAGAGAGAGAGAGAG 1437
Db 754 CGATAATGAGGCTCAATCATTTATCCATGCGCAACACAGATCTCAAGTAGGTACAG 813
Qy 1438 GTCCGACTGCGACCGCGCGCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1497
Db 814 GACCTCATTTGATCCAAAGTCTTTGACCATCTTTCATCAAGAGAGAGAGAGAGAGAG 873
Qy 1498 AGTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1557
Db 874 AAGTCTTTGTCGACAAATCAATGCAATCCATTCGTCCTCCCAATCCCAAGGCTTTCG 933
Qy 1558 ACCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1617
Db 934 ATATTGTTGATCATTCTATGCTCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 993
Qy 1618 CGTCTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1677
Db 994 CGGTTGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1053
Qy 1678 ACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1737
Db 1054 ACAAGTGGTGAACCAACCAAGTATATTTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1113
Qy 1738 ACTTCAGTGGGCGGCTGCTGAGTTCACGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1797
Db 1114 ACTTCAGTGGTGGTATGTTCTTCTGAGTTCACCTCAAAACATTCAGAGAGAGAGAG 1173
Qy 1798 CGTTTCAGGCTTCTCCGACTGGCTTAAT 1826
Db 1174 CTCTCGATTCTCTTCAATTTGGGTTATT 1202

RESULT 5
US-09-295-306-5
; Sequence 5, Application US/09295306
; Patent No. 6198021
; GENERAL INFORMATION:
; APPLICANT: Theodor LANGE et al.
; TITLE OF INVENTION: GA 20-OXIDASE GENE SEQUENCES (As Amended)
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
; STREET: 2033 K Street, N.W., Suite 800

```

CITY: Washington  
 STATE: D.C.  
 COUNTRY: U.S.A.  
 ZIP: 20006  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: MS-DOS  
 SOFTWARE: Wordperfect 5.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/295,306  
 FILING DATE: 08/29/1995  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/553,367  
 FILING DATE: No. 6198021ember 27, 1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Lee Cheng  
 REGISTRATION NUMBER: 40,949  
 REFERENCE/DOCKET NUMBER: 49/DIV-FD4.5M2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-721-8200  
 TELEFAX: 202-721-8250  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 5:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1490 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA to mRNA  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Arabidopsis thaliana  
 IMMEDIATE SOURCE:  
 CLONE: PAT2353  
 US-09-295-306-5

Query Match 10.8%; Score 271; DB 4; Length 1490;

Best Local Similarity 56.5%; Pred. No. 1e-49;  
Matches 593; Conservative 0; Mismatches 435; Indels 21; Gaps 4;

QY 778 TGTTGACGCGCGGTGCTGAGCGGAGAGAGATCCGTCGACGTTTCATATGCGCG 837  
 DB 175 TCCTTAACCTTCTCTTAACTCCATCCCAATCCCAACCAATTCATTGGCCAG 234  
 QY 838 CGAGGAGAGAGCGCGGTGCTGAGAGAGCTGAGGTGGCGCTGATCGAGTGG 897  
 DB 235 ACGAAGAGAAAC---TTCCATTGACATTCAGAGCTCAAGTCCCGCTTCATCGATCT 291  
 QY 898 GGGCGGGCGGAGAGGTGCTGCTGCGGAGTGGGGAGGCTGGAGAGGACG 957  
 DB 292 CAGGCCAAGACTCGACTTTGAAGCTCCTAGAGTCATCGAGAGCTTGACCAACACG 351  
 QY 958 GCTTCTCTGTTGTTAAACAGCGGATCGAGCGCGCTCTGAGAGAGCGCAACGGT 1017  
 DB 352 GCTTCTCTCTGCTGCTCAATCATGCGGTCAGAGTCACTAATAGGAGTCTCACCGTT 411  
 QY 1018 GCATGACGCGCTTCTACGCTGCGCTGGGGAGAGAGCGGCGGCGCGCGCG 1077  
 DB 412 TGATGGAAGTTCTCTGACATGCTCTCGCGGCAACAGAAAGCTCAGAGAAACCG 471  
 QY 1078 GGGAGAGCTCGCGCTACGCCAGCTTTACGGGGCGCTTCCGCTCCAGCTCCGCTGA 1137  
 DB 472 GTGAGAGTTGTGGCTATGCAAGTACCTTCCGCGAGATTCCTCACTAAGCTCCATGA 531  
 QY 1138 AGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1197  
 DB 532 AGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 585  
 QY 1198 GTGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1257

DB 586 AGATTACTTTTCGATACATTAGGACAAAGATTGCG---AGCAGTTTGGAGAGGTATC 642  
 QY 1258 CGCGTACTGCGCAGAGATGAGCGCGCTGTCTGCTGAGAGTGTCTGCGGGGAGA 1317  
 DB 643 AGACTATTGTGAAGCAATGATTTCTATCACTCAAGATCATGAGCTTCGGGCTTAA 702  
 QY 1318 GCGTGGGATGCTGGAGACCGCGCGCTACTTCCCGGAGATCTTCCAGCGCAAGACT 1377  
 DB 703 GTTAGGGTAACCGAG-----ACTATTCCGAGGATTTTTCGAGAGAGAACGATT 753  
 QY 1378 CATCATGCGCTCAACTACTACCGGGGTGCGAGAGCCACTCGACACGCTGGGCAACG 1437  
 DB 754 CAGTATGAGGCTCAATCATTTCTCCATGCGCAACACCAAGATCTCAGTAGTACG 813  
 QY 1438 GTGCGACTGCGACCCACCTCGCTCACTCCATCCACAGGACCACTGCGGCGCTGG 1497  
 DB 814 GACCTCATTTGATCCAAAGTTCTTTGACCATCTTCATCAAGACCATGCAATGCGCTTC 873  
 QY 1498 AGTGTGGGCGAGGCGGTGGCGGCGCAATCCGCGCTCGCGCGCGCGCGCGCGCTG 1557  
 DB 874 AAGTCTTTGTCGACAAATCAATGGAATCAATTCGTCCTCAATCCCAAGGCTTTT 933  
 QY 1558 AGTGTGGGCGACCTTCATGCGCTCTCCAAAGCCAGGTACCGGCTGCTGCGACCGG 1617  
 DB 934 ATATTGGTGAACATTTTATGCTCTATCAAGCGGATATTCAGAGCTGTTGCATAG 993  
 QY 1618 CGTGTCTCAAGACACGCGCGCTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1677  
 DB 994 CGTGTGTAATAGAGAGCGCGGAGAAATCGATGGCGTTTCTGCTGCGGAGAAAG 1053  
 QY 1678 ACAGGTGTGGCGCGCGCGGAGAGGCTGCTGCGAGGACCAACCGGAGGTGTACCGG 1737  
 DB 1054 ACAAGTGTGAACCAACCAAGTATTTTGGAGAGATGAAACAAAGAAATACCTG 1113  
 QY 1738 ACTTCAGTGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1797  
 DB 1114 ACTTCAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1173  
 QY 1798 CGTTCAGGCTTCTCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1826  
 DB 1174 CTCGATTCCTTTCTAATTTGGTTATT 1202

RESULT 6  
 US-09-734-719-5  
 ; Sequence 5, Application US/09734719  
 ; Patent No. 6455675  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Theodor LANGE et al.  
 ; TITLE OF INVENTION: GA 20-OXIDASE GENE SEQUENCES (As Amended)  
 ; NUMBER OF SEQUENCES: 19  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.  
 ; STREET: 2033 K Street, N.W., Suite 800  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: U.S.A.  
 ; ZIP: 20006  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: MS-DOS  
 ; SOFTWARE: Wordperfect 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/734,719  
 ; FILING DATE: 13-Dec-2000  
 ; CLASSIFICATION: <unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 09/295,306  
 ; FILING DATE: April 21, 1999  
 ; APPLICATION NUMBER: 08/553,367  
 ; FILING DATE: No. 6455675ember 27, 1995  
 ; APPLICATION NUMBER: PCT/EP94/01664



```
; ORIGINAL SOURCE:
; ORGANISM: Cucurbita maxima
; IMMEDIATE SOURCE:
; CLONE: Clone pb11
US-08-553-367A-1

Query Match          9.1%; Score 226.4; DB 2; Length 1395;
Best Local Similarity 55.2%; Pred. No. 4.5e-40;
Matches 496; Conservative 0; Mismatches 381; Indels 21; Gaps 2;

QY 928 GGCAGGTGGGGAGGCGTGGGAGGAGCGGCTCTTCTGCTGGTTAAACCAAGCATCG 987
DB 247 GGTGTGTGATGAGGCTGTGAGACACATGGCATATTTTGTGGTGAACCATGAGTGG 306
QY 988 AGGCGGCGTGTGTGAGGAGCGCACCGGTCATGGAGCGCTTCTTACGCTGCCGCTGG 1047
DB 307 ACATAGAAATGATGGCGGCTTTCATGACTGTATGAATGAGTTCTTTACAAATGCCCTTGG 366
QY 1048 GGGAGAGGAGCGGCGGAGCGGCGGCGGAGAGCTGGGCTAGCCAGCAGCTTCA 1107
DB 367 ATGTGAAGCAGAGGCGCTAAGAGGAAGTAGGTAGGATATGAGATATACCAATAGCTTCT 426
QY 1108 CGGGGCGCTTCGCGPCCAAAGTCGCGTGAAGGAGAGCGCTGCTTCCGCTACTCATCGG 1167
DB 427 TTGGAGATTCGCGPCCAAATCTTCCATGGAAGAAACCTTTTCCCTTCGCTGTGGCTG 486
QY 1168 CTGGAGATGAAGAGGCGGCGGAGGCGGCTGGGTGAGTACCTGGTCGGAAGAGCTCGGGGCGG 1227
DB 487 CTCAAAACTCTCCCGGGCTCATGACTATGTTCTTGACACTTTAGGCCCATCATCTTCC 546
QY 1228 AGCAGGCGGCGGCTGGGCGAGGTGACTTCGCTACTGCGACAGATGAGCGCGCTGT 1287
DB 547 ATCATGGG-----AAGCGTATCAAGAGTGGATAGCATTTGAACGAGCTTG 594
QY 1288 OGCTGGAGCTGATGAGGTGCTGGGAGAGAGCTGGGATCTCGGAGAGCGGCGCACT 1347
DB 595 GTACGAAGATTTGGAGCTTTTGGGGCTTAGCCTTGGCATTTCAAGAG-----AAT 645
QY 1348 ACTTCGGCGATCTTTCAGCGGCAACGACTCCATCATGCGGCTCAACTACTACCGGGGT 1407
DB 646 ACTTCGAAGATTTCTTCAGGACAAAGATTCAATATGAGGCTTAATTAATACCCCAACT 705
QY 1408 GCCAGAGGCGACTTCGACACGCTGGGCAACCGCTCGGCACTGGGACCGCCACCTCTCACA 1467
DB 706 GCGACAGCGCAGAGGTGTGTGGGAACTGGCCCTACACACTGATCCCACTCCGTCACAA 765
QY 1468 TCCTCACCAGGACACAGTGGGCGGCTGGAGGTGGGCGGAGGCGGCTGGCGGCGCA 1527
DB 766 TCCCTCACCAGACCGCTGTGAGTGGGCTTCAAGTGTGCTTCAATGATCAATGATATTCAA 825
QY 1528 TCCGCGCTTCGCGCGGCGCTGCTGCTCAACGTCGGGACACCTTTCATGGCGCTCTCA 1587
DB 826 TTCTCCTCAACCCAGAGGCTTTGTGTCATCAACATCGGTGACACTTTCACCTCTCTCA 885
QY 1588 AGCGCAGGTACCGGAGCTGCTGTCACCGGCGGTGCTCAACAGCAGCGGCGCTTCGCGCT 1647
DB 886 ATGGGATTTACAGGCGCTGCATACCGCGCTGTAGTGAATTCATGAAATCCAGAAATAA 945
QY 1648 GCGTGGCTTCTTCCTCTGCGCGGAGATGACACGCTGTGGCGCGCGCGGAGGAGCTGG 1707
DB 946 CATTTGGCTTCTTCTGTGCTTCCATCGCATGACAAAGTGTGAGAGACCGGAGAAATGG 1005
QY 1708 TCGACGACCAACACCGGAGGTGTACCGGAGCTTCACTGTGGCGGCGCTGCTGAGCTTCA 1767
DB 1006 TGGAGAGAGTCCACCAAGAGATCCAGATTAATGAGGCAATGTGCTTGAATGA 1065
QY 1768 GCGAGCGGCTACAGGCGGAGGAGGCGAGCTTCAGGCGCTTCAGGCGCTTCAGGCTTAA 1825
DB 1066 CCGGAGGCGCTTACGAGCTGATTCAGACACTTTTGAAGCGCTTCAAAACCTTGGGCTCA 1123

RESULT 8
US-09-295-306-1
```

```
; Sequence 1, Application US/09295306
; Patent No. 6198021
; GENERAL INFORMATION:
; APPLICANT: Theodor LANGER et al.
; TITLE OF INVENTION: GA-20-OXIDASE GENE SEQUENCES (As Amended)
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Fonack, L.L.P.
; STREET: 2033 K Street, N.W., Suite 800
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/295,306
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/553,367
; FILING DATE: No. 6198021ember 27, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee Cheng
; REGISTRATION NUMBER: 40,949
; REFERENCE/DOCKET NUMBER: 49/DIV-FD4.5M2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-721-8200
; TELEFAX: 202-721-8250
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1395 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Cucurbita maxima
; IMMEDIATE SOURCE:
; CLONE: Clone pb11
US-09-295-306-1

Query Match          9.1%; Score 226.4; DB 4; Length 1395;
Best Local Similarity 55.2%; Pred. No. 4.5e-40;
Matches 496; Conservative 0; Mismatches 381; Indels 21; Gaps 2;

QY 928 GGCAGGTGGGGAGGCGTGGGAGGAGCGGCTCTTCTGCTGGTTAAACCAAGCATCG 987
DB 247 GGTGTGTGATGAGGCTGTGAGACACATGGCATATTTTGTGGTGAACCATGAGTGG 306
QY 988 AGGCGGCGTGTGTGAGGAGCGCACCGGTCATGGAGCGCTTCTTACGCTGCCGCTGG 1047
DB 307 ACATAGAAATGATGGCGGCTTTCATGACTGTATGAATGAGTTCTTTACAAATGCCCTTGG 366
QY 1048 GGGAGAGGAGCGGCGGAGCGGCGGCGGAGAGCTGGGCTAGCCAGCAGCTTCA 1107
DB 367 ATGTGAAGCAGAGGCGCTAAGAGGAAGTAGGTAGGATATGAGATATACCAATAGCTTCT 426
QY 1108 CGGGGCGCTTCGCGPCCAAAGTCGCGTGAAGGAGAGCGCTGCTTCCGCTACTCATCGG 1167
DB 427 TTGGAGATTCGCGPCCAAATCTTCCATGGAAGAAACCTTTTCCCTTCGCTGTGGCTG 486
QY 1168 CTGGAGATGAAGAGGCGGCGGAGGCGGCTGGGTGAGTACCTGGTCGGAAGAGCTCGGGGCGG 1227
DB 487 CTCAAAACTCTCCCGGGCTCATGACTATGTTCTTGACACTTTAGGCCCATCATCTTCC 546
QY 1228 AGCAGGCGGCGGCTGGGCGAGGTGACTTCGCTACTGCGACAGATGAGCGCGCTGT 1287
```

[illegible]

## RESULT 9

US-09-734-719-1  
Sequence 1, Application US/09734719  
Patent No. 6455675  
GENERAL INFORMATION:  
APPLICANT: Theodor LANGE et al.  
TITLE OF INVENTION: GA 20-OXIDASE GENE SEQUENCES (AS Amended)  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.  
STREET: 2033 K Street, N.W., Suite 800  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09734, 719  
FILING DATE: 13-Dec-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/295,306  
FILING DATE: April 21, 1999  
APPLICATION NUMBER: 08/553,367  
FILING DATE: No. 6455675/ender 27, 1995  
APPLICATION NUMBER: PCT/EP94/01664  
FILING DATE: May 24, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee Cheng











GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2003, 05:23:30 ; Search time 137.899 Seconds  
(without alignments)  
15433.627 Million cell updates/sec

Title: US-09-702-134-7212\_COPY\_28000\_30500

Perfect score: 2501  
Sequence: 1 atgtccaagataaaattggc.....ctatcaatgtgttccaat 2501

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications, NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/US06\_PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	226.4	9.1	1161	10	US-09-371-307-77
2	153.2	6.1	1077	10	US-09-371-307-8
3	129	5.2	811	10	US-09-371-307-71
4	84.8	3.4	12733	9	US-10-032-393-47
5	84.8	3.4	12739	9	US-10-032-393-8
6	81.2	3.2	266	10	US-09-923-876-1829
7	80	3.2	768	9	US-09-938-842A-812
8	74.8	3.0	1179	9	US-09-938-842A-1511
9	73.2	2.9	276	10	US-09-294-093B-5184
10	73.2	2.9	1050	9	US-09-938-842A-2672
11	73	2.9	446	10	US-09-864-761-20699
12	71.6	2.9	1075	10	US-09-864-761-19241
13	71.6	2.9	1403	10	US-09-864-761-2513
14	71.4	2.9	1030	10	US-09-371-307-5
15	71.4	2.9	1852	10	US-09-969-852-4
c 16	68.2	2.7	1635	10	US-09-864-761-20241
c 17	68.2	2.7	1973	10	US-09-864-761-3471
c 18	68.2	2.7	2283	9	US-09-938-842A-1691
c 19	68.2	2.7	2283	10	US-09-887-576-645

c 20	66.8	2.7	2561	9	US-09-976-740-48
c 21	66.8	2.7	2561	12	US-10-023-529-48
c 22	66.8	2.7	2561	12	US-10-023-523-48
c 23	66.4	2.7	401	10	US-09-864-761-3936
c 24	65	2.6	1064	10	US-09-804-682-29
c 25	64.6	2.6	1083	9	US-09-938-842A-2267
c 26	64	2.6	1077	9	US-09-938-842A-1529
c 27	63.2	2.5	426	10	US-09-371-307-70
c 28	62.8	2.5	830	10	US-09-864-761-19531
c 29	62.4	2.5	598	10	US-09-770-149-991
c 30	61.6	2.5	479	10	US-09-864-761-768
c 31	59.8	2.4	4257	9	US-09-825-288A-1
c 32	59	2.4	520	9	US-10-184-644-332
c 33	59	2.4	4282	10	US-09-778-927A-16
c 34	58.4	2.3	14800	10	US-09-954-456-1601
c 35	58.2	2.3	15720	9	US-10-025-380-1058
c 36	58.2	2.3	15720	10	US-09-922-217-1058
c 37	58.2	2.3	15720	10	US-09-833-263-1058
c 38	57.8	2.3	1456	10	US-09-764-877-3391
c 39	57.6	2.3	1062	9	US-09-938-842A-2398
c 40	57	2.3	245	10	US-09-923-876-2158
c 41	56.8	2.3	267	10	US-09-864-761-27984
c 42	56.8	2.3	1174	10	US-09-864-761-11284
c 43	56.8	2.3	3809	12	US-10-001-870-68
c 44	56.4	2.3	4826	10	US-09-772-304A-1
c 45	56.2	2.2	1062	9	US-09-938-842A-1357

ALIGNMENTS

RESULT 1  
US-09-371-307-77  
; Sequence 77, Application US/09371307A  
; Patent No. US20020053095A1  
; GENERAL INFORMATION:  
; APPLICANT: Brown, Sherri M.  
; APPLICANT: Heck, Gregory R.  
; APPLICANT: Piller, Kenneth J.  
; APPLICANT: Kishore, Ganesh M.  
; APPLICANT: Elich, Tedd D.  
; APPLICANT: Logusch, Eugene W.  
; APPLICANT: Rao, Sudabathula  
; APPLICANT: Ream, Joel E.  
; APPLICANT: Logusch, Sherry J.  
; TITLE OF INVENTION: Methods for controlling gibberellin levels  
; FILE REFERENCE: MOET:216  
; CURRENT APPLICATION NUMBER: US/09/371,307A  
; CURRENT FILING DATE: 1999-08-10  
; NUMBER OF SEQ ID NOS: 89  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 77  
; LENGTH: 1161  
; TYPE: DNA  
; ORGANISM: Cucurbita maxima  
US-09-371-307-77

Query Match 9.1%; Score 226.4; DB 10; Length 1161;  
Best Local Similarity 55.2%; Pred. No. 1.4e-50;  
Matches 496; Conservative 0; Mismatches 381; Indels 21; Gaps 2;

QY	928	GGCAGGTCGGGGGCGCGCGGAGGCGCGGCGCGATGCGCGCTCTTCACGCTGCCTGG 1047
DB	233	GGCTGCTGGATGAGGCTTGTAGCAACATGCGCATATTTTGTGGTGAACCATGGAGTGG 292
QY	988	AGGCGCGCTGCTGGAGGAGGCGCGCGGCGCGATGCGCGCTCTTCACGCTGCCTGG 1047
DB	293	ACATAGAAATAGTGGGCGCGCTGTTATGACTGTATGAATGAGTCTTTACATGCGCTTGG 352
QY	1048	GGGAGAAACGCGCGCGAGCGCGCGGAGAGCTGCGGCTACGCCAGCGCTTCA 1107
DB	353	ATGTGACGAGAGGCGCTAGAGAGGAGGTAGGTGAGAGTTATGGATATACCAATAGCTTCT 412



## RESULT 3

US-09-371-307-71

; Sequence 71, Application US/09371307A

; Patent No. US20020053095A1

; GENERAL INFORMATION:

; APPLICANT: Brown, Sherri M.

; APPLICANT: Heck, Gregory R.

; APPLICANT: Pillier, Kenneth J.

; APPLICANT: Kishore, Ganesh M.

; APPLICANT: Elich, Tadd D.

; APPLICANT: Logusch, Eugene W.

; APPLICANT: Rao, Sudabathula

; APPLICANT: Ream, Joel E.

; APPLICANT: Logusch, Sherry J.

; TITLE OF INVENTION: Methods for controlling gibberellin levels

; FILE REFERENCE: MOBT-216

; CURRENT APPLICATION NUMBER: US/09/371,307A

; CURRENT FILING DATE: 1999-08-10

; NUMBER OF SEQ ID NOS: 89

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 71

; LENGTH: 811

; TYPE: DNA

; ORGANISM: Zea mays

US-09-371-307-71

Query Match

Best Local Similarity 5.2%; Score 129; DB 10; Length 811;

Matches 278; Conservative 0; Mismatches 205; Indels 6; Gaps 2;

QY 1274 GATGAGCCGCTGCTGCTGAGCTGATGAGGTGCTCGGGGAGAGCGTGGCATGCTGG 1333

Db 22 GATGAGCTGCTGCTGCTGAGCTGATGAGGTGCTCGGGGAGAGCGTGGCATGCTGG 81

QY 1334 AGACGGCGCCACTACTTCGCGGCGATCTTCCAGCGCAACGACTCCATCATGCGCTCAA 1393

Db 82 GGACGGGTGCTGGCGCGCTGTGTGCGCGCGCGGACGACGCTGCTGCGGGTGA 141

QY 1394 CTACTACCGCGGTGCCAGAGGCCACTCGACACGCT---GGGACCGGCTCCGCACTGGCA 1450

Db 142 CCACTACCGCGCGCGCGCTCAACCCAGCGCTCACGSGGTTCGGGAGCACACCGA 201

QY 1451 CCCCACCTCGCTACCACTTCCTCACAGAGCAACGCTCGCGCGCTGAG---GTGTGGGC 1507

Db 202 CCGCGAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 261

QY 1508 GGAGGGGGGTGGCGCGCCATCGCGCGCTCGCGCGCGCGCTCGCTCAAGCTCGGGA 1567

Db 262 GGAGCGGCTGGGCTCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 321

QY 1568 CACCTTCATGCGGTCTCTCAACGCGAGGTACCGAGCTGCTCGACCGCGCGCTGCTCAA 1627

Db 322 CACCTTCATGCGGTCTCTCAACGCGAGGTACCGAGCTGCTCGACCGCGCGCTGCTCAA 381

QY 1628 CACGACGGCGCTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1687

Db 382 CACGAGAGTCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 441

QY 1688 GCGCCCGCGGAGGAGTGTTCGACGACCAACCGAGGCTGTACCGGACTTCACCTG 1747

Db 442 GCGCCCGCTTCGCGAGCTCTCTGGCGACGCGCGCGCGCGCGCGCGCGCGCGCGCG 501

QY 1748 GCGGCGCT 1756

Db 502 GAGCGAGT 510

## RESULT 4

US-10-032-393-47

; Sequence 47, Application US/10032393

; Publication No. US20030027286A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Wall, Daniel

; APPLICANT: Gross, Molly

; TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE

; FILE REFERENCE: ELITRA.010A

; CURRENT APPLICATION NUMBER: US/10/032,393

; CURRENT FILING DATE: 2001-12-21

; PRIOR APPLICATION NUMBER: 60/259,434

; PRIOR FILING DATE: 2000-12-27

; PRIOR APPLICATION NUMBER: 09/948,993

; PRIOR FILING DATE: 2001-09-06

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; NUMBER OF SEQ ID NOS: 68

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 47

; LENGTH: 12733

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Vector pPEPF14

US-10-032-393-47

Query Match 3.4%; Score 84.8; DB 9; Length 12733;

Best Local Similarity 45.9%; Pred. NO. 9.2e-12;

Matches 290; Conservative 0; Mismatches 342; Indels 0; Gaps 0;

QY 711 TTAATGATTGATAGCTAGATTATCAACAATTATGAGCATGGTGTGCAGCAGGACAG 770

Db 4460 TATATGTTTATGGCTGACCTTATCTTCAAGTGGGGGGGGGGGGGGGGGGGG 4519

QY 771 GAGGTGGTGTTCGACGGCGGCTGCTGAGCGGGCAGACGAGATCCCGTCGCGAGTCA 830

Db 4520 GGG 4579

QY 831 TGGCGGGGGGGGAGACCCCGGGTGGTGGGGTGGAGAGCTGGAGGTGGCGTGCATC 890

Db 4580 GGG 4639

QY 891 GAGCTGGGGGGGGGGGGGGGAGGTGCTGCTGGTGGTCCGGGAGAGTGGGGGAGGCTGCAG 950

Db 4640 GGG 4699

QY 951 AGCAGCGCTTCTTCCTGCTGTTTAAACAGGCTATCAGAGCGCGGCTGCTGAGAGAGCG 1010

Db 4700 GGG 4759

QY 1011 CACCGGTGCTAGGACGCTTCTTACGCTGCGGCTGGGGGAGAGAGAGCGGCGAGCGG 1070

Db 4760 GGG 4819

QY 1071 CGCGCGGGGGAGAGTGGGCTACGCCAGAGCTTTCACGGGCGCTTTCGCTCCAGCTG 1130

Db 4820 GGG 4879

QY 1131 CCCTGAGAGAGAGCTGCTGCTTCCGCTACTCATCGCTGAGATGAAGAGGCGAGAG 1190

Db 4880 GGG 4939

QY 1191 GCGTGGGTGAGTACTGCTGCGAAAGCTCGGGGCGGAGCAGCGGCGGCTGGCGAG 1250

Db 4940 GGG 4999

QY 1251 GTCTACTGCGCTACTGCTGACAGATGAGCCGCTGCTGCTGAGAGTGTGAGTGTCTC 1310

Db 5000 GGG 5059

QY 1311 GGGGAGAGCTGGCATCTGTCGGAGACCGGCG 1342

Db 5060 GGG 5091







QY 1450 ACCCCACCTCGTCACTCACTCCACAGGACACAGCTGGCGGCTGGAGTGTGGCGG 1509  
 Db 96 ACATGAGCAGCTCAGCTCCCTCTGTCGCCACAGAGTGCAGGGCTCCAGTCTTAAG 155  
 QY 1510 AGGGGGGTTGGCGGCTCACTCCCTCGCCCGGGGCTCGTCTCAAGTGGCGACA 1569  
 Db 156 ATGTCATGAGTACGAGGCAAGTACGTGGCCGAGCAGCTATGCTCCATCGGCGATC 215  
 QY 1570 CTTTCATGCGCTCTCCACGCGCAGGTACCGAGCTGCTCCACGGGCGGTCTCAAC 1629  
 Db 216 AGATCGAGATTTTCAGCAACAGGGGCGATACAGGCGGTGCTGCACCGTACGAGGTGAACA 275

RESULT 10  
 US-09-938-842A-2672  
 ; Sequence 2672, Application US/09938842A  
 ; Patent No. US20020160378A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Harper, Jeff  
 ; APPLICANT: Kreps, Joel  
 ; APPLICANT: Wang, Xun  
 ; APPLICANT: Zhu, Tong  
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
 ; FILE REFERENCE: SAME, AND METHODS OF USE  
 ; CURRENT APPLICATION NUMBER: US/09/938,842A  
 ; PRIOR FILING DATE: 2001-08-24  
 ; PRIOR APPLICATION NUMBER: US 60/227,866  
 ; PRIOR FILING DATE: 2000-08-24  
 ; PRIOR APPLICATION NUMBER: US 60/264,647  
 ; PRIOR FILING DATE: 2001-01-16  
 ; PRIOR APPLICATION NUMBER: US 60/300,111  
 ; PRIOR FILING DATE: 2001-06-22  
 ; NUMBER OF SEQ ID NOS: 5379  
 ; SEQ ID NO 2672  
 ; LENGTH: 1050  
 ; TYPE: DNA  
 ; ORGANISM: Arabidopsis thaliana  
 ; US-09-938-842A-2672

Query Match 2.9%; Score 73.2; DB 9; Length 1050;  
 Best Local Similarity 50.9%; Pred. No. 2.3e-09;  
 Matches 174; Conservative 0; Mismatches 168; Indels 0; Gaps 0;

QY 1382 CATGGCTCACTACTACCGGCTGCCAGAGCCACTGACACGCTGGCGACCGGTCC 1441  
 Db 600 CATGGCTTAACTACTATCGCGCTGTCACAGAACCGAGCTAACTTACGAGCTCCCGG 659  
 QY 1442 GCACCTGGACCCACCTCCCTCACCATCTCCACAGGACCGCTGGCGGCTGGAGGT 1501  
 Db 660 ACATAAGACCCACCGCTTATCAGCTCTTCTTCAAGACCAAGTCTCTGGTTTCAAGT 719  
 QY 1502 GTGGCGGAGGGGCGGTGCGGCGCATCGGCTCGCGGCGGCTGCTGCAACGT 1561  
 Db 720 CTTTAAGGNTAATAATGGTGCCTGTAGTCCATTCACACACTTTCATGTCATAT 779  
 QY 1562 CGCGACACCTTATCGGCTCTCCACGCGAGTACCGAGCTGCTGCGACCGGCGGT 1621  
 Db 780 CGCGACCAATGTCAGGTATAGCAATGATAAATACAGAGTGTGCTCCATAGAGCGGT 839  
 QY 1622 CGTCAACAGCAGCGGCTCGCGGCTCGCTGCTGCTCTTCTGCGCGGAGTGGACAC 1681  
 Db 840 AGTAARACCGAGAACGAGCGGTATTCGATTCGAGCTTTCTATTCCTTCACAGATGC 899  
 QY 1682 GTGGTGGCGCGCGGAGGAGTGTGTCGACGACCCACCC 1723  
 Db 900 ACTGATTGTTCCACGACAGAGCTGCTCAATGACAGAGACTC 941

RESULT 11  
 US-09-864-761-20699  
 ; Sequence 20699, Application US/09864761

Patent No. US20020048763A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Penn, Sharon G.  
 ; APPLICANT: Rank, David R.  
 ; APPLICANT: Hanzel, David K.  
 ; APPLICANT: Chen, Wensheng  
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL F  
 ; FILE REFERENCE: Aemica-X-1  
 ; CURRENT APPLICATION NUMBER: US/09/864,761  
 ; CURRENT FILING DATE: 2001-05-23  
 ; PRIOR APPLICATION NUMBER: US 60/180,312  
 ; PRIOR FILING DATE: 2000-02-04  
 ; PRIOR APPLICATION NUMBER: US 60/207,456  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: US 09/632,366  
 ; PRIOR FILING DATE: 2000-08-03  
 ; PRIOR APPLICATION NUMBER: GB 24263.5  
 ; PRIOR FILING DATE: 2000-10-04  
 ; PRIOR APPLICATION NUMBER: US 60/236,359  
 ; PRIOR FILING DATE: 2000-09-27  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: US 60/234,687  
 ; PRIOR FILING DATE: 2000-09-21  
 ; PRIOR APPLICATION NUMBER: US 09/608,408  
 ; PRIOR FILING DATE: 2000-06-30  
 ; PRIOR APPLICATION NUMBER: US 09/774,203  
 ; PRIOR FILING DATE: 2001-01-29  
 ; NUMBER OF SEQ ID NOS: 49117  
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
 ; SEQ ID NO 20699  
 ; LENGTH: 446  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: MAP TO AC007249.2  
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 10  
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 14  
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 13  
 ; OTHER INFORMATION: EXPRESSED IN HELI00, SIGNAL = 12  
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 13  
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 13  
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 12  
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 13  
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 12  
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 9.7  
 ; OTHER INFORMATION: NT HIT: AL161539.2, EVALUATE 3.70e-00  
 ; US-09-864-761-20699

Query Match 2.9%; Score 73; DB 10; Length 446;  
 Best Local Similarity 50.7%; Pred. No. 1.5e-09;  
 Matches 175; Conservative 0; Mismatches 170; Indels 0; Gaps 0;  
 QY 738 CAAATATGACATGTGTGTCGACGAGGAGGTGTGTCAGCGCGGCTGCTG 757



Qy	751	TGGTGGTCGACGACAGGAGCAGAGAGTGGTGTTCGACGCGGGCGGTCTGACGGGCGACACGG	810
Dd	431	TGGTCATGGTGGTGGTATGGTGGTGGTATGGTATGGTGGTGGTATGGTGGTGGTGGTGGTGGTGG	490

QY	1382	CATGCGCTCAACTACTACCGCGCGTGCAGAGGCCACTCGACAGCTGTGGCCACCGCTCC	1441
Db	615	CTTGCATTCACTCTTACCGCACTTGTCCGGATCCGGATCGGGTCATGGGTCTGCCGC	674
QY	1442	GCATCGGACCCCACTCGTCTACCAATCTCCACAGGACACGCTCGCGCGCTCGGAGGT	1501

```
Db 675 CCACACCGGACTCCACCCCTTCTCACAAATCCCTTACCAAAACAAACATAAGCGGGTTGAGGT 734
QY 1502 GT---GEGCGGAGGCGGTGGGCGCCCATCCGCGCCCTCGCCCGCGGCGGTGCTGTGTCAA 1558
Db 735 TCACCGAAAAGCGCGGGTGGGTGACGGTGGCACACAGTCCCGAAGGGCTTGTGATCAA 794
QY 1559 GCTCGGGGACACCTTCATGCGGCTCTCCACACGCCAGGTACCGCAGCTGCTGACCGGGC 1618
Db 795 TGTAGCGACCTCCCTCCACATATTTGCGAACGGGTGTACCGGAGTGTGCTCCACCGGGT 854
QY 1619 GGTCTGCAACAGACGCGGCGCTCCGCGCTCGCTGCTGCTCTCTCTGCGCGGAGATGA 1678
Db 855 CTTAGTGACCGGATCCAGCAAGGCTTTCAGTGGGTATTTATGTGGGCCCCCTCCGAA 914
QY 1679 CAGCGTGTGCGCCGCGGAGGAGTGTGTCAGCAGCACCAACCCGAGGGTGTACCCGGA 1738
Db 915 TGTGGAGATATGTCACATGCGAAGCTAGTGGGCCCAATAAGCTCCCTTTATAAGGC 974
QY 1739 CTTACAGTGGCGGCGCTGCTGGAC 1763
Db 975 AGTCACTTGAATGAGTACCTTGGC 999

RESULT 15
US-09-969-852-4
; Sequence 4, Application US/09969852
; Patent No. US20020137211A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Tianyan
; APPLICANT: Liu, Huiyan
; APPLICANT: Li, Wei
; APPLICANT: Zhao, Libin
; TITLE OF INVENTION: A METHOD FOR ESTABLISHING AN EXPRESSION SYSTEM OF SPIDER DRAGLINE
; FILE REFERENCE: LIU-65
; CURRENT APPLICATION NUMBER: US/09/969,852
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: CN01106406.4
; PRIOR FILING DATE: 2001-01-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1852
; TYPE: DNA
; ORGANISM: Nephila clavipes
US-09-969-852-4

Query Match 2.9%; Score 71.4; DB 10; Length 1852;
Best Local Similarity 46.5%; Pred. No. 1e-08;
Matches 271; Conservative 0; Mismatches 306; Indels 6; Gaps 1;

QY 761 GCAGGACGAGGAGTGTTCGACGCGGGGTGCTGACGGGGAGACGGAGATCCCGTC 820
Db 151 GCTGGAGGAGGAGTATAGTGGACAAGGTGCAGGTGCAGCAGCAGCAGCAGCAGCAGCTGGA 210
QY 821 GCAGTTCATATATGCCGCGGAGAGAGCCCGGCTCGGTGGCGGTGGAGGAGCTGGAGGT 880
Db 211 GGTTCGGACAAGAGGACTAGTGGACAGGTGCTGGACAGGAGCTGGAGCAGCCGCT 270
QY 881 GGCCTGATCGAGTGGGGGGGGGGGAGAGCTGTCGGTGGTCCGCGAGGTGGGGGA 940
Db 271 GCACGAGCTGCTGTCGCGACAGGAGGATATGGAGTCTCGGAAGCCAAAGGTGCAGGA 330
QY 941 GSCGTGAGAGGACGCGCTTTCTGCTGTTTACCAAGGCATCGAGGCGCGCTGCT 1000
Db 331 CAGGTGATTCAGTGGACAAGGGCAGGTGTCAGCAGCAGCAGCAGCTGGAGGTGCCGGA 390
QY 1001 GGAGGAGCGCACCGGTGCATGAGCGCTCTTT-----CACGCTCCCGCTGGGGGAGAA 1054
Db 391 CAGGAGGATATGAGGTCTTGGABGCCAAGGTCCAGGACGAGGTGGATTAGTGGACAA 450
QY 1055 GCAGGGGCGCAGCGCGCGGGGAGAGCTGCGGCTACGCCACGACGCTTCACGGGGCG 1114
```

Search completed: March 26, 2003, 14:09:04  
Job time : 182.899 secs

```
Db 451 GGTGAGGTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTGT 510
QY 1115 CTTCCGCTCCAAAGCTGCCGTGGAAGGAGACGCTGCTGCTCCGCTACTCATFCGGCTGGAGA 1174
Db 511 CTTGCTGGACAAGGTGCCGACACAAGGTGGCTATGAGGAGCTTGGAAAGCCAAGGTGCCCGGA 570
QY 1175 TGAAGAGGCGAGGAGGCGCTGGGTGAGTACCTGTCGCGAAGCTCGGGCGGAGCACGG 1234
Db 571 CGAGGAGGATTAGGTGGACAAGGTGCAGGTGCAGCAGCAGCAGCAGCAGCTGGAGGTGCC 630
QY 1235 GCGCGGCTGGCGGAGGTGCTACTCGCGCTACTGCCACGAGATGAGCCGCTGCTGCTGGA 1294
Db 631 GGCACAGGAGGAGCTAGTGGACAAAGGTGCTGGACAAAGGCTGGAGCAGCCCTGCACGA 690
QY 1295 GCTGATGAGAGTGTTCGGGAGGAGCCCTGGGCATCTGTCGGAGAC 1337
Db 691 GCTGCTGGTCCGGGACAAGGAGGATATGGAGGTCTTTGGAAGCC 733
```

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: March 26, 2003, 04:18:30 ; Search time 9640.7 Seconds  
(without alignments)  
15096.740 Million cell updates/sec  
Title: US-09-702-134-7212\_COPY\_62300\_67300  
Perfect score: 5001  
Sequence: 1 aggggacactacacttcca.....tgccgcgagtgccgccg 5001  
Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*  
29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rtd.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Match	Length	DB	ID	Description
c	1	5001	100.0	145999	2	AC096690	AC096690 Oryza sat
c	2	2033.2	40.7	207376	2	AC104433	AC104433 Oryza sat
c	3	938.8	18.8	207376	2	AC104433	AC104433 Oryza sat
c	4	315.2	6.3	131487	2	AP004117	AP004117 Oryza sat
c	5	274.8	5.5	159290	2	AP004813	AP004813 Oryza sat
c	6	263	5.3	96360	2	AP003867	AP003867 Oryza sat
c	7	263	5.3	134787	2	AP005490	AP005490 Oryza sat
c	8	261.4	5.2	127498	2	AP004759	AP004759 Oryza sat
c	9	231.8	4.6	145999	2	AC096690	AC096690 Oryza sat
c	10	231	4.6	121752	2	AP004130	AP004130 Oryza sat
c	11	231	4.6	176224	2	AP005163	AP005163 Oryza sat
c	12	217.8	4.4	152835	8	AP004224	AP004224 Oryza sat
c	13	216.6	4.3	163704	2	AP004849	AP004849 Oryza sat
c	14	215.8	4.3	126637	2	AC118980	AC118980 Oryza sat
c	15	215.4	4.3	138906	2	AP003946	AP003946 Oryza sat
c	16	215.4	4.3	143160	2	AP004796	AP004796 Oryza sat
c	17	215.4	4.3	148272	2	AP004156	AP004156 Oryza sat
c	18	215.4	4.3	169000	2	AP004991	AP004991 Oryza sat
c	19	215	4.3	121996	2	AP003952	AP003952 Oryza sat
c	20	213.4	4.3	148533	2	AP005449	AP005449 Oryza sat
c	21	213.4	4.3	172368	2	AP003712	AP003712 Oryza sat
c	22	213.4	4.3	175743	2	AP003579	AP003579 Oryza sat
c	23	213	4.3	178959	2	AP004809	AP004809 Oryza sat
c	24	211.4	4.2	111333	2	AC120534	AC120534 Oryza sat
c	25	210.2	4.2	151997	2	OSJN01010	OSJN01010 Oryza sat
c	26	210.2	4.2	156266	8	AP002542	AP002542 Oryza sat
c	27	209.6	4.2	131768	2	AP005476	AP005476 Oryza sat
c	28	208.4	4.2	142544	2	OSJN00286	OSJN00286 Oryza sat
c	29	208.2	4.2	172386	2	OSJN00223	OSJN00223 Oryza sat
c	30	208	4.2	151343	2	AP005003	AP005003 Oryza sat
c	31	207.6	4.2	143515	8	AP003252	AP003252 Oryza sat
c	32	205.8	4.1	124047	2	AC112209	AC112209 Oryza sat
c	33	205.2	4.1	144583	8	AP003262	AP003262 Oryza sat
c	34	205.2	4.1	160925	8	AP004332	AP004332 Oryza sat
c	35	204.2	4.1	114184	2	AP003984	AP003984 Oryza sat
c	36	204.2	4.1	153428	2	AP005535	AP005535 Oryza sat
c	37	203	4.1	176580	2	AP003634	AP003634 Oryza sat
c	38	202.6	4.1	130263	8	OSJN00065	OSJN00065 Oryza sat
c	39	202.4	4.0	114062	2	OSJN00267	OSJN00267 Oryza sat
c	40	202	4.0	176349	8	AP002092	AP002092 Oryza sat
c	41	201.6	4.0	137152	2	AP004564	AP004564 Oryza sat
c	42	201.6	4.0	171970	2	AP004764	AP004764 Oryza sat
c	43	201.4	4.0	159420	2	AP005459	AP005459 Oryza sat
c	44	199.8	4.0	145739	8	OSJN00157	OSJN00157 Oryza sat
c	45	199.6	4.0	122882	2	AC093088	AC093088 Oryza sat

ALIGNMENTS

RESULT 1  
AC096690/c  
LOCUS  
DEFINITION  
Oryza sativa chromosome 3 clone OSJNba0059G06, \*\*\* SEQUENCING IN  
PROGRESS \*\*\*  
ACCESSION  
AC096690  
VERSION  
AC096690.4 GI:20503077  
KEYWORDS  
HTG; HTGS\_PHASE2.  
SOURCE  
Oryza sativa.  
ORGANISM  
Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.  
REFERENCE  
1 (bases 1 to 145999)  
AUTHORS  
Buell,C.R., Yuan,Q., Ouyang,S., Liu,J., Tallon,L., Gansberger,K.,





QY 1441 ACAGGGAACTGCGCGAGCGAAGAGGCGCTTAGCGAAGGACATCCCGGAGCAGCGCG 1500  
Db 79679 ACAGGAACTGCGCGAGCGAAGAGGCGCTTAGCGAAGGACATCCCGGAGCAGCGCG 79620  
QY 1501 CAGGTGGGTGAGTCCGATCCCTTCTGAGAGGCGCTGCGGCTCTCCCATCGGAGC 1560  
Db 79619 CAGGTGGGTGAGTCCGATCCCTTCTGAGAGGCGCTGCGGCTCTCCCATCGGAGC 79560  
QY 1561 TGGGGGAGAGGCGGAGGATGGAGGTTGGGGGCGCACTCCGGGGGAAACCATCGCGTCC 1620  
Db 79559 TGGGGGAGAGGCGGAGGATGGAGGTTGGGGGCGCACTCCGGGGGAAACCATCGCGTCC 79500  
QY 1621 CAGTGGGACCCACACAGCCGCCGCCGCCCATGTTGCGATGCGCTCCCGAGTGAAC 1680  
Db 79499 CAGTGGGACCCACACAGCCGCCGCCGCCCATGTTGCGATGCGCTCCCGAGTGAAC 79440  
QY 1681 ACCACCTTGTTCAGCGAGGAGGAGACACAGCCGCCGCCGCCCGGGAAGCAGCAGCAAC 1740  
Db 79439 ACCACCTTGTTCAGCGAGGAGGAGACACAGCCGCCGCCGCCCGGGAAGCAGCAGCAAC 79380  
QY 1741 AGCAGCAGCATCGCGCGATTCGCTGCTGCTACCTCTCACTACTCTACTGCTAGTCAATTTAT 1800  
Db 79379 AGCAGCAGCATCGCGCGATTCGCTGCTGCTACCTCTCACTACTCTACTGCTAGTCAATTTAT 79320  
QY 1801 GTTTGGGACAAATATATTTGAATCTACTCTCTCACTACTCTACTGCTAGTCAATTTAT 1860  
Db 79319 GTTTGGGACAAATATATTTGAATCTACTCTCTCACTACTCTACTGCTAGTCAATTTAT 79260  
QY 1861 AATCGAGCTAGCGCATGCTCTGTTGAGTTGTTGACCCCAACCAACCAACCAACCA 1920  
Db 79259 AATCGAGCTAGCGCATGCTCTGTTGAGTTGTTGACCCCAACCAACCAACCAACCA 79200  
QY 1921 GAAGGAAACACACACCCACCATGCTGCGCTGCTCCCTTTCTTTCTTTCTTTCTTT 1980  
Db 79199 GAAGGAAACACACATCCACCATGCTGCGCTGCTCCCTTTCTTTCTTTCTTTCTTT 79140  
QY 1981 TTTTGGGTTTGTGTTGCGGAGTATAGCAGGAAGAAAGTAGCAGTGGGCTAGTT 2040  
Db 79139 TTTTGGGTTTGTGTTGCGGAGTATAGCAGGAAGAAAGTAGCAGTGGGCTAGTT 79080  
QY 2041 TGTGAGTGGCTGGGCTCTTATTCCTAGGGCTAGCGCCAGGTGCATCGGAGGAGCGA 2100  
Db 79079 TGTGAGTGGCTGGGCTCTTATTCCTAGGGCTAGCGCCAGGTGCATCGGAGGAGCGA 79020  
QY 2101 GCCAGGCTGCTGCGCTGCTCGGACTACGATACGATGAGTGGAGTGGAGTCAAG 2160  
Db 79019 GCCAGGCTGCTGCGCTGCTCGGACTACGATACGATGAGTGGAGTGGAGTCAAG 78960  
QY 2161 TCGTCGCCATTCGCGCGCCCGATCAGCTTTCATCTGSAATTCAAACCATCTTCTCT 2220  
Db 78959 TCGTCGCCATTCGCGCGCCCGATCAGCTTTCATCTGSAATTCAAACCATCTTCTCT 78900  
QY 2221 CCTCCGATCTTGCATTAACACATTCACCTAGAGGTGATCAAGAGGTAGCTAGGC 2280  
Db 78899 CCTCCGATCTTGCATTAACACATTCACCTAGAGGTGATCAAGAGGTAGCTAGGC 78840  
QY 2281 CGGAACATCATCATTCAGGTAGCGTCACTAAATTTTACTGAAATACACTACTACTAG 2340  
Db 78839 CGGAACATCATCATTCAGGTAGCGTCACTAAATTTTACTGAAATACACTACTACTAG 78780  
QY 2341 TAGTACCACACCATAGTGTCTGGGCAACCAAGCTGATCTTGGTTATCGGCTGTGAT 2400  
Db 78779 TAGTACCACACCATAGTGTCTGGGCAACCAAGCTGATCTTGGTTATCGGCTGTGAT 78720  
QY 2401 TTTCTGAGTACAGGCTTAATATCTGATATCAAGCTGATCAATCACTTATACAGTCA 2460  
Db 78719 TTTCTGAGTACAGGCTTAATATCTGATATCAAGCTGATCAATCACTTATACAGTCA 78660  
QY 2461 AGAATGACTTGCCTATGCAATAGAGGAGGAGCAGCATGACGAGCTAGCAGAGTACAA 2520  
Db 78659 AGAATGACTTGCCTATGCAATAGAGGAGGAGCAGCATGACGAGCTAGCAGAGTACAA 78600

QY 2521 TGTAAACAACCTCGAGACCTCCATGAATCCGAGATTGCCCTCACACGAGAAACGACCGAC 2580  
Db 78599 TGTAAACAACCTCGAGACCTCCATGAATCCGAGATTGCCCTCACACGAGAAACGACCGAC 78540  
QY 2581 TTGCTTTTATCCTTGGTCCCTTCACTCATCTCCCTAGCTCGCTCCCTCCCTCTCTTCT 2640  
Db 78539 TTGCTTTTATCCTTGGTCCCTTCACTCATCTCCCTAGCTCGCTCCCTCTCTTCTTCT 78480  
QY 2641 CCTCCCCACGACCAACAAATGTTGACCTCGCTGCGAGTTTACCTTTCTACTTGCCTGAT 2700  
Db 78479 CCTCCCCACGACCAACAAATGTTGACCTCGCTGCGAGTTTACCTTTCTACTTGCCTGAT 78420  
QY 2701 TTCACCGATTCCAAACAGCTACATTAACCTCGTTAGAGTCCACATGCCATTTGATGTGTGT 2760  
Db 78419 TTCACCGATTCCAAACAGCTACATTAACCTCGTTAGAGTCCACATGCCATTTGATGTGTGT 78360  
QY 2761 GGGGAAAGGGAAGAGGATGTTGGGATGTCATCTGTTGGGATGATGGCATCGTCTCTC 2820  
Db 78359 GGGGAAAGGGAAGAGGATGTTGGGATGTCATCTGTTGGGATGATGGCATCGTCTCTC 78300  
QY 2821 CAGCAAGGAGGATGGGATGTTGATTGGAGTGAAGAGGATTTGATTGGAGTGAAGGATTTG 2880  
Db 78259 CAGCAAGGAGGATGGGATTTGATTGGAGTGAAGAGGATTTGATTGGAGTGAAGGATTTG 78240  
QY 2881 GGAGGAGCAAAAGTGACGACGAGAGAAATGAGTGGAGAGGAAGAGAGAGAAACAGA 2940  
Db 78239 GGAGGAGCAAAAGTGACGACGAGAGAAATGAGTGGAGAGGAAGAGAGAGAAACAGA 78180  
QY 2941 TATGGAAATAAGAAATAAATCTAGGGTGTCTTGTGTTAAGTGTATCCCATTTCTACGGG 3000  
Db 78179 TATGGAAATAAGAAATAAATCTAGGGTGTCTTGTGTTAAGTGTATCCCATTTCTACGGG 78120  
QY 3001 ATCTAGGGTGTATTTGTTAGTAGTGTACTAGTATCTCTAATTCACGGAGTCCCATCCA 3060  
Db 78119 ATCTAGGGTGTATTTGTTAGTAGTGTACTAGTATCTCTAATTCACGGAGTCCCATCCA 78060  
QY 3061 CTAGTTCCTGTAATAAATAATATGCGCTTTTGACCGTATCGTACTAGTAGTGAATAATG 3120  
Db 78059 CTAGTTCCTGTAATAAATAATATGCGCTTTTGACCGTATCGTACTAGTAGTGAATAATG 78000  
QY 3121 ACCATTGCTGATGAGAGAGAGAAATAGAGTGAATGATGAGATAAAATATAGGA 3180  
Db 77999 ACCATTGCTGATGAGAGAGAGAAATAGAGTGAATGATGAGATAAAATATAGGA 77940  
QY 3181 AAATGTTGTTCCATACCATCGAAAGTTCCTGAGTTTTCCTTATCCATCGAAATTTG 3240  
Db 77939 AAATGTTGTTCCATACCATCGAAAGTTCCTGAGTTTTCCTTATACCATCGAAATTTG 77880  
QY 3241 GTGCATTTAATAACCATCAAAGTTTTCCTGTTAACTCTCTATACCACTGTTTAT 3300  
Db 77879 GTGCATTTAATAACCATCAAAGTTTTCCTGTTAACTCTCTATACCACTGTTTAT 77820  
QY 3301 TCCGTCAGTTTTTTTTTCTTTTCAATATTTGGCTGGAAGACAAACAGGATGACTGTAT 3360  
Db 77819 TCCGTCAGTTTTTTTTTCTTTTCAATATTTGGCTGGAAGACAAACAGGATGACTGTAT 77760  
QY 3361 TGCCTTCTCTTCACTCACTAAATCTGCCCCAAATTTTCTGTTTTTCTCTTTTGGCTG 3420  
Db 77759 TGCCTTCTCTTCACTCACTAAATCTGCCCCAAATTTTCTGTTTTTCTCTTTGGCTG 77700  
QY 3421 ACATGGGACATCGCTCCATCGATGGGATCGGAGCAACAAAGAGCAGGAGCAGCAGGCC 3480  
Db 77699 ACATGGGACATCGCTCCATCGATGGGATCGGAGCAACAAAGAGCAGGAGCAGCAGGCC 77640  
QY 3481 GGAGGAACATTTGTAATGAGACTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3540  
Db 77639 GGAGGAACATTTGTAATGAGACTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 77580  
QY 3541 CGCGCGCGCGGAGGTGACGCGCAGCAGGAGCGGCTGTCGCCCTGCTCTCTCTATC 3600  
Db 77579 CGCGCGCGCGGAGGTGACGCGCAGCAGGAGCGGCTGTCGCCCTGCTCTCTCTATC 77520  
QY 3601 TTTCCCGCGCATTCGCCAAATCTACGTCCTCTGCTACTGATCTCTTCCCTACCGGTGAG 3660





FEATURES	Location/Qualifiers	source
*	25643	25601: gap of unknown length
*	25692	31250: contig of 5559 bp in length
*	31251	31299: gap of unknown length
*	31300	34474: contig of 3175 bp in length
*	34475	34523: gap of unknown length
*	34524	37120: contig of 2597 bp in length
*	37121	37169: gap of unknown length
*	37170	39229: contig of 2060 bp in length
*	39230	39278: gap of unknown length
*	39279	41558: contig of 2280 bp in length
*	41559	41607: gap of unknown length
*	41608	47008: contig of 5401 bp in length
*	47009	47057: gap of unknown length
*	47058	49299: contig of 2242 bp in length
*	49300	49348: gap of unknown length
*	49349	52145: contig of 2797 bp in length
*	52146	52194: gap of unknown length
*	52195	70888: contig of 18694 bp in length
*	70889	70937: gap of unknown length
*	70938	82563: contig of 11626 bp in length
*	82564	82612: gap of unknown length
*	82613	89680: contig of 7068 bp in length
*	89681	89729: gap of unknown length
*	89730	94211: contig of 4482 bp in length
*	94212	94259: gap of unknown length
*	94260	96345: contig of 2056 bp in length
*	96346	96363: gap of unknown length
*	96364	101734: contig of 5371 bp in length
*	101735	101782: gap of unknown length
*	101783	127272: contig of 25490 bp in length
*	127273	127320: gap of unknown length
*	127321	139823: contig of 12503 bp in length
*	139824	139871: gap of unknown length
*	139872	161256: contig of 21395 bp in length
*	161257	161304: gap of unknown length
*	161305	163387: contig of 2083 bp in length
*	163388	163435: gap of unknown length
*	163436	165397: contig of 2162 bp in length
*	165398	165645: gap of unknown length
*	165646	167865: contig of 2220 bp in length
*	167866	167913: gap of unknown length
*	167914	171872: contig of 3959 bp in length
*	171873	171920: gap of unknown length
*	171921	174530: contig of 2610 bp in length
*	174531	174578: gap of unknown length
*	174579	176604: contig of 2026 bp in length
*	176605	176652: gap of unknown length
*	176653	178799: contig of 2147 bp in length
*	178800	178847: gap of unknown length
*	178848	180996: contig of 2149 bp in length
*	180997	181044: gap of unknown length
*	181045	185204: contig of 4160 bp in length
*	185205	185252: gap of unknown length
*	185253	195354: contig of 10102 bp in length
*	195355	195402: gap of unknown length
*	195403	198616: contig of 3214 bp in length
*	198617	198654: gap of unknown length
*	198655	207376: contig of 8712 bp in length.

QY	1295	CCCCCTGGTACCAAGAGACAGACGCCGATGCGCGCGCGCGGTCTCTCGAGGGCGACCCCTGGAGC	1355
Db	39281		
QY	1355	GGCGACACAGATCGGAGTAGAGTGGTGGCCCTTGGCCCACTGTCGGCCATCTCGTGGCCGC	1414
Db	39221	GGCGACACAGATCGGAGTAGAGTGGTGGCCCTTGGCCCACTGTCGGCCATCTCGTGGCCGC	39162
QY	1415	CGACGGCCAGGGGACGAGCGCGATGACAGGAACTGGCCGAGCGAGAGGAGGCGGTAG	1474
Db	39161	CGACGGCCAGGGGACGAGCGCGATGACAGGAACTGGCCGAGCGAGAGGAGGCGGTAG	39102
QY	1475	CGAAGGACATCCCGGGACGACGCCGCGAGCTGCGGTTGGAGTCGATGCCGTGTGGAGCG	1534
Db	39101	CGAAGGACATCCCGGGACGACGCCGCGAGCTGCGGTTGGAGTCGATGCCGTGTGGAGCG	39042
QY	1535	GCCTGTGGCCCTCTCCCATCGGAGCTGCGGGGAGAGCGCGGAGGATGAGAGGGTGTGGGG	1594
Db	39041	GCCTGTGGCCCTCTCTCCCATCGGAGCTGCGGGGAGAGCGCGGAGGATGAGAGGGTGTGGGG	38982
QY	1595	CGCACTCCGGGGGAACCATCGCTGCCAGTGGGACCCACACCGCCCCCGCCCGCCCA	1654
Db	38981	CGCACTCCGGGGGACCATCGCGTCCAGTGGGACCCACACCGCCCCCGCCCGCCCA	38922
QY	1655	TGTTCGACTGCCCTCCGAGATGAACACCACTTGTTCGACGAGGAGGACACAGCG	1714
Db	38921	TGTTCGACTGCCCTCCGAGATGAACACCACTTGTTCGACGAGGAGGACACAGCG	38862
QY	1715	CGCCGCGCGCGGAGCAGCAGCAGCAGCATCCGCGCATTCGCTGCTGCTACC	1774
Db	38861	CGCCGCGCGCGGAGCAGCAGCAGCAGCATCCGCGCATTCGCTGCTGCTACC	38802
QY	1775	GTTCGGCTTGGGGCCTGGGGTGTGGGACAAATAATTTGAAATCTACTCTCTCT	1834
Db	38801	GTTCGGCTTGGGGCCTGGGGTGTGGGACAAATAATTTGAAATCTACTCTCTCT	38742
QY	1835	CACCTACTACTCTAGTCAATTTATAANGAGCTACGGCAGTGTCTGTTGAGTTGTT	1894
Db	38741	CACCTACTACTCTAGTCAATTTATAANGAGCTACGGCAGTGTCTGTTGAGTTGTT	38682
QY	1895	TGACCCACCAACCAACCAACCAAGAAAGAAACACACATCCACCCATGTCGTGGCGCTG	1954
Db	38681	TGACCCACCAACCAACCAACCAAGAAAGAAAGAAACACATCCACCCATGTCGTGGCGCTG	38622
QY	1955	CTCTCCCTCTTTCTTTCTTTCTTTTCTTTTGGGTTTGTTCGCGAGTATAGCAGGA	2014
Db	38621	CTCTCTCCCTCTTTCTTTCTTTTCTTTTGGGTTTGTTCGCGAGTATAGCAGGA	38562
QY	2015	AGAAAGTAGTACGATGGGCGTAGTTTGTTCGATGGCGCTGGGCTCCCTTATCGAGGCT	2074
Db	38562	AGAAAGTAGTACGATGGGCGTAGTTTGTTCGATGGCGCTGGGCTCCCTTATCGAGGCT	38502
QY	2075	ACGGCCCGAGTGCATCGGAGGAGCGACGAGCTCTTCGCTGCTCCGGAATCAGT	2134
Db	38502	ACGGCCCGAGTGCATCGGAGGAGCGACGAGCTCTTCGCTGCTCCGGAATCAGT	38442
QY	2135	ATACTGGACTGGACTGAGGTACAGTTCGTCGACATTCCTGCGCCCGCCGATCACTCTTC	2194
Db	38442	ATACTGGACTGGACTGAGGTACAGTTCGTCGACATTCCTGCGCCCGCCGATCACTCTTC	38382
QY	2195	CATCTGGATTCAAAACCATCTTCTTCCTCCGATCTTCGCCATAACACATTTCACTCAG	2254
Db	38382	CATCTGGATTCAAAACCATCTTCTTCCTCCGATCTTCGCCATAACACATTTCACTCAG	38322
QY	2255	CAAGTGCATCAAGGTTAGCTAGCCCGGACATTCATTCAGCGTACCGTCAACTAA	2314
Db	38322	CAAGTGCATCAAAAGTTAGCTAGCCCGGACATTCATTCAGCGTACCGTCAACTAA	38262
QY	2315	ATTTTACTGAAATACACTACTACTAGTAGTACCACACCTAGGTGTCTGGGACACAGCC	2374
Db	38262	ATTTTACTGAAATACACTACTACTAGTAGTACCACACCTAGGTGTCTGGGACACAGCC	38202
QY	2375	TGATACCTGTGTATCGGCTGTGATTTCTTTCGAGTACACAGGCTAAATCTTGGTATCAG	2434

```

Db 38202 TGTACTTGTGGTTATCGGCTGTGATTTCTGACGTACCGGGCTAATACCTTGATGAC 38143
QY 2435 CGGTGATTAATCACCTTATACCACTAGCAATGACTGTGCTATGCAACAATGAAGAGGACA 2494
Db 38142 CGGTGATTAATCACCTTATACCACTAGCAATGACTGTGCTATGCAACAATGAAGAGGACA 38083
QY 2495 GCATGACGAGTACGACAGTACATGTAACAACCTCGAGACCTCCATGAACCTCCGAGA 2554
Db 38082 GCATGACGAGTACGACAGTACATGTAACAACCTCGAGACCTCCATGAACCTCCGAGA 38023
QY 2555 TTGCGCTCACACCCAGAACCGGACTGCTTTTATCCTTGGTGCCTTTGACCTCATCTC 2614
Db 38022 TTGCGCTCACACCCAGAACCGGACTGCTTTTATCCTTGGTGCCTTTGACCTCATCTC 37963
QY 2615 CCTCAGCTCGGCTCCCTCTTGTCCCTCCAGCCAGCAGCAAAATGTGACCTCGTG 2674
Db 37962 CCTCAGCTCGGCTCCCTCTTGTCCCTCCAGCCAGCAGCAAAATGTGACCTCGTG 37903
QY 2675 TCGCAGTTACCTTCTACTTCCGATTTCCCGATTTCCCGATTTCCAGAGCTACAACTCGTTAG 2734
Db 37902 TCGCAGTTACCTTCTACTTCCGATTTCCCGATTTCCAGAGCTACAACTCGTTAG 37843
QY 2735 AGTCACATCCGATTTGATGTGTGTGGGAAGGGAAGAGGATGTGGATGTCTATGCT 2794
Db 37842 AGTCACATCCGATTTGATGTGTGTGGGAAGGGAAGAGGATGTGGATGTCTATGCT 37783
QY 2795 GTTGGGATGATGCGGATCGCTCTCCAGCAAAAGGAGTGGATTCATTGGAGTGAAG 2854
Db 37782 GTTGGGATGATGCGGATCGCTCTCCAGCAAAAGGAGTGGATTCATTGGAGTGAAG 37723
QY 2855 AGGATGATGAGTGAAGATTTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGT 2914
Db 37722 AGGATGATGAGTGAAGATTTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGT 37663
QY 2915 GGAGAGGAGAAAGAGGAGAAACAGATATGAGATAGAAATAATCTAGGCTGTTGT 2974
Db 37662 GGAGAGGAGAAAGAGGAGAAACAGATATGAGATAGAAATAATCTAGGCTGTTGT 37603
QY 2975 TGTAAAGTGTATCCCTTCTAGGAGTCTAGGCTGTTTATGTTAAGTAGTGTACTAG 3034
Db 37602 TGTAAAGTGTATCCCTTCTAGGAGTCTAGGCTGTTTATGTTAAGTAGTGTACTAG 37543
QY 3035 TATCTTAATTCACGGGACTCCATCCACTAGTTCCTGTAATGAAATAATGCGCTCTTTCAC 3094
Db 37542 TATCTTAATTCACGGGACTCCATCCACTAGTTCCTGTAATGAAATAATGCGCTCTTTCAC 37483
QY 3095 CGTATCGTACTAGTGAATGACCAATGACCAATGCGTATGAGAGAGAAATAGGATG 3154
Db 37482 CGTATCGTACTAGTGAATGACCAATGACCAATGCGTATGAGAGAGAAATAGGATG 37423
QY 3155 AATGATGATGAGATAAATAATGGAATAATGGTTCCATACCATCGAAAGTTCCTGAGTT 3214
Db 37422 AATGATGATGAGATAAATAATGGAATAATGGTTCCATACCATCGAAAGTTCCTGAGTT 37363
QY 3215 TTGCTTTAACCATCGAAATAATGTTGCTGCTGCTTAAATACCATTCAAAGTTTTCCTGTT 3274
Db 37362 TTGCTTTAACCATCGAAATAATGTTGCTGCTGCTTAAATACCATTCAAAGTTTTCCTGTT 37303
QY 3275 AACTCTCTATACCACTGTTTATTTTCCTGCTGCTGCTTTCCTTTCATATTTGGCT 3334
Db 37302 AACTCTCTATACCACTGTTTATTTTCCTGCTGCTGCTTTCCTTTCATATTTGGCT 37243
QY 3335 GGAAGCAAAACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3394
Db 37242 GGAAGCAAAACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 37183
QY 3395 TTTTCTCGTTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCT 3454
Db 37182 ATTTCTCGTTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCT 37123
QY 3455 AGCA 3458

```

```

Db 37122 NNCA 37119

RESULT 3
AC104433
LOCUS
DEFINITION
Oryza sativa chromosome 3 clone OJ1754_E06, *** SEQUENCING IN
PROGRESS ***, 34 unordered pieces.
AC104433
AC104433
AC104433.3 GI:19111667
HTG; HTGS_PHASE1.
SOURCE
Oryza sativa
ORGANISM
Oryza sativa
REFERENCE
AUTHORS
Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Eukaryota; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 207376)
Buell, C.R., Yuan, Q., Ouyang, S., Liu, J., Tallon, L., Gansberger, K.,
Kim, M., Overton, I.L., Bera, J., Tsitrin, T., Krol, M., Jarrahi, B.,
Jin, S., Koo, H., Zismann, V., Hsiao, J., Blunt, S., Vanaken, S.,
Utterbach, T., Feldblyum, T., Yang, Q., Haas, B., Suh, B., Peterson, J.,
Quackenbush, J., White, O., Salzberg, S. and Fraser, C.
Oryza sativa ssp. japonica cv. Nipponbare OJ1754_E06 BAC genomic
sequence
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 207376)
Buell, R.
TITLE
Direct Submission
JOURNAL
Submitted (11-DEC-2001) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
REFERENCE
3 (bases 1 to 207376)
Buell, R.
TITLE
Direct Submission
JOURNAL
Submitted (02-APR-2002) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
COMMENT
On Mar 5, 2002 this sequence version replaced gi:18767486.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 34 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 10102: contig of 10102 bp in length
* 10103 10151: gap of unknown length
* 10152 13105: contig of 2954 bp in length
* 13106 13154: gap of unknown length
* 13155 20540: contig of 7386 bp in length
* 20541 20589: gap of unknown length
* 20590 22982: contig of 2393 bp in length
* 22983 23031: gap of unknown length
* 23032 25642: contig of 2611 bp in length
* 25643 25691: gap of unknown length
* 25692 31250: contig of 5559 bp in length
* 31251 31299: gap of unknown length
* 31300 34474: contig of 3175 bp in length
* 34475 34523: gap of unknown length
* 34524 37120: contig of 2597 bp in length
* 37121 37169: gap of unknown length
* 37170 39229: contig of 2060 bp in length
* 39230 39278: gap of unknown length
* 39279 41558: contig of 2280 bp in length
* 41559 41607: gap of unknown length
* 41608 47008: contig of 5401 bp in length
* 47009 47057: gap of unknown length
* 47058 49299: contig of 2242 bp in length
* 49300 49348: gap of unknown length
* 49349 52145: contig of 2797 bp in length
* 52146 52194: gap of unknown length
* 52195 70888: contig of 18694 bp in length
* 70889 70937: gap of unknown length
* 70938 82563: contig of 11626 bp in length

```

QY	301	C	TAAAGGTGCAGATGGCCATCTTCAAGTTGCGAACACCCCTTGTCATCACAATAATCTCACGTTT	360
Ds	167226	C	TAAAGGTGCAGATGGCCATCTTCAAGTTGCGAACACCCCTTGTCATCACAATAATCTCACGTTT	167285
QY	361	G	GAAGTTTGATCCCTTCTCGAGCTTCCTTACTACTTCAGTATATTGCCAAGCCCTGAT	420
Ds	167286	G	GAAGTTTGATCCCTTCTCGAGCTTCCTTACTACTTCAGTATATTGCCAAGCCCTGAT	167345
QY	421	G	GTAGTCCCTACCTGAAACACCACAGAGAATGGGGATTCGCTACTGTAATATGATGATA	480
Ds	167346	G	GTAGTCCCTACCTGAAACACCACAGAGAATGGGGATTCGCTACTGTAATATGATGATA	167405
QY	481	T	PACAGCGAGAGAGTCTAATTTACAGCGATTTTCAGCAATATCCCAATATAGTTTGTCTTG	540
Ds	167406	T	PACAGCGAGAGAGTCTAATTTACAGCGATTTTCAGCAATATCCCAATATAGTTTGTCTTG	167465
QY	541	C	AATTCGCTGGGTGGTTGAATGAGGAGAAATATGCAATGGTGCATAGTAAGATGACAG	600
Ds	167466	C	AATTCGCTGGGTGGTTGAATGAGGAGAAATATGCAATGGTGCATAGTAAGATGACAG	167525
QY	601	T	AGTTGTGCATGTAGACCTTGGACATATATATAGCCTATTCTATACAGATATATATAGGGA	660
Ds	167526	T	AGTTGTGCATGTAGACCTTGGACATATATATAGCCTATTCTATACAGATATATATAGGGA	167585
QY	661	T	TGTGTGCATATCATGCATGATTGAGCTGAGGTGGCGGTGGCAACATATAATGCA	720
Ds	167586	T	TGTGTGCATATCATGCATGATTGAGCTGAGGTGGCGGTGGCAACATATAATGCA	167645
QY	721	T	TACCATCATCTCAGACTTTGAAAACCAACAATCTAAGTAAGATGGAATGCTTAAACT	780
Ds	167646	T	TACCATCATCTCAGACTTTGAAAACCAACAATCTAAGTAAGATGGAATGCTTAAACT	167705
QY	781	A	CAATGGAAATTTATTTTGGCATGCTACCTCGACGAATGTGATGGTATGTTATTACATTA	840
Ds	167706	A	CAATGGAAATTTATTTTGGCATGCTACCTCGACGAATGTGATGGTATGTTATTACATTA	167765
QY	841	G	GAGAGACACCAACAGACACAGCAGAGGAATAGTAGCTTGGAAATATAAGTATGTTGG	900
Ds	167766	G	GAGAGACACCAACAGACACAGCAGAGGAATAGTAGCTTGGAAATATAAGTATGTTGG	167825
QY	901	A	GTAGTATAGCTGAAATCTTGTATGCACAGCCTATTAAACA	940
Ds	167826	A	GTAGTATAGCTGAAATCTTGTATGCACAGCCTATTAAACA	167865
<hr/>				
RESULT 4				
AP004117				
LOCUS	AP004117	131487 bp	DNA	linear HTG 21-MAR-2000
DEFINITION	Oryza sativa (japonica cultivar-group) chromosome 2 clone OJ134_F06, ** SEQUENCING IN PROGRESS ***, in ordered pieces.			
ACCESSION	AP004117			
VERSION	AP004117.1	GI:15375103		
KEYWORDS	HTG; HYGS PHAGE2.			
SOURCE	Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA, clone:OJ134_F06			
ORGANISM	Oryza sativa (japonica cultivar-group)			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.			
REFERENCE	1			
AUTHORS	Sasaki, T., Matsumoto, T. and Yamamoto, K.			
TITLE	Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, BAC clone:OJ134_F06			
JOURNAL	Published Only in Database (2001)			
REFERENCE	2 (bases 1 to 131487)			
AUTHORS	Sasaki, T., Matsumoto, T. and Yamamoto, K.			
TITLE	Direct Submission			
JOURNAL	Submitted (29-AUG-2001) takuji Sasaki, National Institute of Agricultural Sciences, Rice Genome Research Program; Kamondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan (E-mail:Tsasakienias@affrc.go.jp, URL:http://irg.dna.affrc.go.jp/Tel:81-298-38-7441, Fax:81-298-38-7468)			
	The nucleotide sequence of this BAC clone was generated by			

combining Monsanto and RGP-Japan sequencing data.

NOTE: It currently consists of 1 conigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter. This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.

- \* NOTE: This is a 'working draft' sequence.
- \* This sequence will be replaced
- \* by the finished sequence as soon as it is available and
- \* the accession number will be preserved.

FEATURES	source	Location/Qualifiers
		i. .131487
		/organism="Oryza sativa (japonica cultivar-group)"
		/cultivar="Nipponbare"
		/db_xref="taxon:39947"
		/chromosome="2"
		/clone="OJ1134_F06"
BASE COUNT	38675 a 26762 c 26756 g 39195 t	99 others
ORIGIN		
Query Match	6.3%;	Score 315.2; DB 2; Length 131487;
Best Local Similarity	70.1%;	Pred. No. 2.4e-42;
Matches 597;	Conservative 0;	Mismatches 155; Indels 97; Gaps 9;
QY	3219	TTTATACATCGAAATTTGTGGTCGACTTATATACCATCAAGCTTTTCGCTAACT 3278
Db	44240	TATATCCATCGAAATTTGTGGTCGACTTATATACCATCAAGCTTTTCGCTAACT 44299
QY	3279	CTCTATACCAACTGTTTATTTCGGTCAGTTT-TTTTTCTTTTCAATTTGGTGGG 3337
Db	44300	TTCTATACCAACTGTTTATTTCGGTTAGTTTGTCTTTTCAATTTGGCTGGA 44359
QY	3338	AGACAACAGGATATGACTGTATTCGCTTCCTTCACATCATTAATTCGCCCAATTT 3397
Db	44360	AAACAACAGGATATGACTGTATTCGCTTCCTTCACATCATTAATTCGCCCA 44415
QY	3398	TCTCGTTTTCCTTTTGGCTGACATGGGCACATCGTCCATCATGGATGGGAGCG 3457
Db	44416	-----TTTTTCTTCTTTGGCCGACATAGGCAATCGCTTCATCAATGGGATGGGAGC 44470
QY	3458	AACAAGAGCA- GGCAGCACAGCGCGGAGGAACATTGTGAATAGAGGACTCGACAGCAC 3515
Db	44471	AACAAGAGCAGCGCAACACAGCGCGGAGGAACATTGTGAGGAGGAGACTCGACAGCAC 44530
QY	3516	CGAGCAGCGGAGTGGTTAGACGCGCGCGCGGAGACTCGACGCGCGACAGGAGCG 3575
Db	44531	CAGACACCGGAGTGGTTAGACGCGCGCGGAGGAGCAGCGCGCGAGGAGCG 44590
QY	3576	-----GCTGTGCGCCCTGCTTCTCTCTATCTTCTCCCGCCCATTCGCG 3618
Db	44591	TGAGCTAGTTCCCAACCATTTTCACTGATCTCAAGCACCATTCATCCCATCCGCG 44650
QY	3619	AATCTACAGTCCGCTGCTACTGATCTCTTCTACGCGTGAAGTGGTCCCAACCATTTCA 3678
Db	44651	TAAATCTGCGCGCGGTGGCCCAATTTGCTCTCAAGTTAGAGCTCTCTTTGTCAATGA 44710
QY	3679	ACTCTCAAGCACCATTTCA-----TCCCATCAATCCGGGTAATCTCG 3719
Db	44711	ACATCATATCGAGCTCAAGCTTCTTCCTGTAGCAGCCCGCTCCAGAGCGCGCTCAGG 44770
QY	3720	CGCGCGGTGGCCCAATTTGC-----CTCAAGCTTCTTCCAAAGCGCGGCCCT 3765
Db	44771	TGCTGTCGGGTTTTCGCCCACTTTTGTGCGAGGTGACCGGCGCGCCCGGCCCT 44830
QY	3766	CCAGGAATTCGAGCTGCGGTCAACATCTGTCTCTCCCGGCGCAA-CCCTTTCCTCACAA 3824
Db	44831	CCAGGAATTCGAGCTGCGGTCAACATCTGTCTCTCCCGGCGCAAACCCCTTTCCTCGCA 44890
QY	3825	GCGGATGGCTG-SCCAGCTCCGACAGGTGCGGTGGCAAGGAGTGTGATGAGCTGGCG 3883
Db	44891	GTGATGGCTGTGGCGGCTCCGACAGGTGCGGTGGCAGGAGTGTGATGAGCTGGCG 44950

QY	3884	CGATCGTGTGTTGAAGAAGATGCTGCTTTTGTATCTGGATCGTGGGGACCGGAGACAAGG	3943
Db	44951	GCA-----ATGCGTCGGGGACCGGACAAGG	44977
QY	3944	GCAACCAATGTCATGTTTAAACACAGATCCACTAAAGTTTGACGGTGTGTAATAACT	4003
Db	44978	GCAAGTAATGTCGTTTAAACACAGATCCACTCAAGTTTCACGGTGTGTAATAAAGG	45037
QY	4004	AACGGAGTAAAT	4015
Db	45038	CCCGTTTGATT	45049
RESULT 5			
AP004813/c			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
BASE COUNT			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 384;			
QY	1444	GGGAACGCGGAGCGGAGGAGGCGGTTACGAGAGGACATCCGGGACCGACCGCGAG	1503
Db	27070	GGCACTGGGCGGCTGGAAGAGGGGGGAGCGCGCGGTTGAGCCTGGGAGGGAGGAGAGG	27011
QY	1504	GTGCGGTTTGAGTGCATGCGGTTGTGGAGCGGCTCGTGGCGCTCTCTCCATCGAGATCG	1563
Db	27010	GGGCGGTTTGAGTGCATGCGGTTGTGGAGCGGCTCGTGGCGCTCTCTCCATCGAGATCG	26951
QY	1564	GGGAGAGCGGAGGATGAGGAGGCGTGGGGGCGCACTCCGGG-GGAACATGCGGTCCTCA	1622





Db 134194 -----AGTTGGGAGGACCAAGGTGGAGGATGCAGAGAATGAAGTTCGAGAGAAAGG 134246

QY 2926 AAGAGGAGAACAGATATGGAGATAA--GAAATAAATCTAGSGTGTTGTGTTAAAGTT 2983  
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 134247 AGGAGAAATAGTCTTAGGAGAAAGAAACCGAAGGATGTTGGGGTGTTCAATTATAAGTA 134306

QY 2984 GTATCCCATCTCAGGGAT 3002  
||||| || |||||

Db 134307 GTATCCTATCTCATPGGGAT 134325

RESULT 8  
AP004759 127498 bp DNA linear HTG 21-MAR-2002  
LOCUS Oriza sativa (japonica cultivar-group) chromosome 8 clone P0670E08.  
DEFINITION \*\*\* SEQUENCING IN PROGRESS \*\*\*, in ordered pieces.

ACCESSION AP004759  
VERSION AP004759.1 GI:18656405  
KEYWORDS HTG; HTGS PHASE2.  
SOURCE Oriza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,  
clone:P0670E08.

ORGANISM Oriza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartioideae; Oryzeae; Oryza.

REFERENCE 1  
AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.  
TITLE Oriza sativa nipponbare(GR3) genomic DNA, chromosome 8, PAC  
clone:P0670E08  
JOURNAL Published Only in Database (2002)

REFERENCE 2 (bases 1 to 127498)  
AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.  
TITLE Direct Submission  
JOURNAL Submitted (13-FEB-2002) Takuji Sasaki, National Institute of  
Agrobiological Sciences, Rice Genome Research Program; Kamondai  
2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rtp.dna.affrc.go.jp/,  
Tel:81-298-38-7441, Fax:81-298-38-7468)  
NOTE: It currently consists of 1 contigs. Gaps between the contigs  
are represented as runs of N. The order of the pieces is believed  
to be correct as given, however the sizes of the gaps between them  
are based on estimates that have provided by the submitter. This  
sequence will be replaced by the finished sequence as soon as it is  
available and the accession number will be preserved.  
\* THIS sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.

FEATURES  
source  
location/Qualifiers  
1..127498  
/organism="Oriza sativa (japonica cultivar-group)"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/chromosomes="8"  
/clone="P0670E08"

BASE COUNT 35025 a 29255 c 28860 g 34022 t 336 others

ORIGIN  
  
Query Match 5.2%; Score 261.4; DB 2; Length 127498;  
Best Local Similarity 68.9%; Pred. No. 2.2e-33;  
Matches 468; Conservative 0; Mismatches 161; Indels 50; Gaps 6

QY 2351 ACCTAGGTGTCGGCACCAAGCCTGACTACTGTGGGTATATCGCTGTGATTTCTGCAGT 2410  
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 76966 ATUUTGGTATCTGCTACCAGGCGTGACTTTGGGTATCATCTCGTATCCCCGAGT 77025

QY 2411 ACAAGCCTATATCTGGTATFATCGGCTGGAAATCACTTATACCACTCAAG----- 2462  
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 77026 ACTATGCCGTATCTTGTTATCAGCTGTGATATCTCGACGTTTCACACTGATATCTGA 77085

QY 2463 -----AATGACTTGCCTATGCAATGAAATGAGAGAGAGCATGACGATGACGACA 2513  
TACTAGTCTAGTAACTTACCTACTATATGAAGAGAGAGACACAAACGATGACGAGG 77145

\* are represented as runs of  $N$ . The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submittor.

*	1	3902:	contig of 3902 bp in length
*	2	4005:	gap of unknown length
*	3	4006:	gap of unknown length
*	4	71267:	contig of 67265 bp in length
*	5	71268:	gap of unknown length
*	6	94679:	contig of 23312 bp in length
*	7	94680:	gap of unknown length
*	8	94780:	gap of unknown length
*	9	142220:	contig of 27441 bp in length
*	10	142221:	gap of unknown length
*	11	142222:	gap of unknown length
*	12	142223:	gap of unknown length
*	13	145999:	contig of 23679 bp in length
*	14	145999:	contig of 23679 bp in length

FEATURES	SOURCE
1. <i>General</i>	
2. <i>Specific</i>	
3. <i>Other</i>	
4. <i>Other</i>	
5. <i>Other</i>	
6. <i>Other</i>	
7. <i>Other</i>	
8. <i>Other</i>	
9. <i>Other</i>	
10. <i>Other</i>	
11. <i>Other</i>	
12. <i>Other</i>	
13. <i>Other</i>	
14. <i>Other</i>	
15. <i>Other</i>	
16. <i>Other</i>	
17. <i>Other</i>	
18. <i>Other</i>	
19. <i>Other</i>	
20. <i>Other</i>	
21. <i>Other</i>	
22. <i>Other</i>	
23. <i>Other</i>	
24. <i>Other</i>	
25. <i>Other</i>	
26. <i>Other</i>	
27. <i>Other</i>	
28. <i>Other</i>	
29. <i>Other</i>	
30. <i>Other</i>	
31. <i>Other</i>	
32. <i>Other</i>	
33. <i>Other</i>	
34. <i>Other</i>	
35. <i>Other</i>	
36. <i>Other</i>	
37. <i>Other</i>	
38. <i>Other</i>	
39. <i>Other</i>	
40. <i>Other</i>	
41. <i>Other</i>	
42. <i>Other</i>	
43. <i>Other</i>	
44. <i>Other</i>	
45. <i>Other</i>	
46. <i>Other</i>	
47. <i>Other</i>	
48. <i>Other</i>	
49. <i>Other</i>	
50. <i>Other</i>	
51. <i>Other</i>	
52. <i>Other</i>	
53. <i>Other</i>	
54. <i>Other</i>	
55. <i>Other</i>	
56. <i>Other</i>	
57. <i>Other</i>	
58. <i>Other</i>	
59. <i>Other</i>	
60. <i>Other</i>	
61. <i>Other</i>	
62. <i>Other</i>	
63. <i>Other</i>	
64. <i>Other</i>	
65. <i>Other</i>	
66. <i>Other</i>	
67. <i>Other</i>	
68. <i>Other</i>	
69. <i>Other</i>	
70. <i>Other</i>	
71. <i>Other</i>	
72. <i>Other</i>	
73. <i>Other</i>	
74. <i>Other</i>	
75. <i>Other</i>	
76. <i>Other</i>	
77. <i>Other</i>	
78. <i>Other</i>	
79. <i>Other</i>	
80. <i>Other</i>	
81. <i>Other</i>	
82. <i>Other</i>	
83. <i>Other</i>	
84. <i>Other</i>	
85. <i>Other</i>	
86. <i>Other</i>	
87. <i>Other</i>	
88. <i>Other</i>	
89. <i>Other</i>	
90. <i>Other</i>	
91. <i>Other</i>	
92. <i>Other</i>	
93. <i>Other</i>	
94. <i>Other</i>	
95. <i>Other</i>	
96. <i>Other</i>	
97. <i>Other</i>	
98. <i>Other</i>	
99. <i>Other</i>	
100. <i>Other</i>	

```

/organism="Oryza sativa"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/chromosome="3"
/clone="OSJNBa0059G06"
/note="japonica cultivar

```

BASE COUNT	40847 a	32003 c	31335 g	41412 t	402 others
ORIGIN	/note- japonica cultivar group				

```
Query Match      4.6%; Score 231.8; DB 2; Length 145999;
Best Local Similarity 64.7%; Pred. No. 1.8e-28;
Matches 385; Conservative 0; Mismatches 192; Indels 18;
```

QY	1183	ACCTGGATGAGGAGAGGTGTGGATGCGCGAGATCGCGCGCAGAGTTCGCGACGAGCATG	1242
Db	75694	ACCTGGATACACGACAGCTCCGCGCAAGCCAGCTCTCGCGACGATCGCGACCATGCC	75753
QY	1243	GCACATCCGGGGGGGCTACTCGTTGGCTTCGCGCCACAGCAGCGTCTCGCTCTCCCTCTGG	1302
Db	75754	TCCATCTTCGGGGCGTACACCTCCGCGTCTCCGCGCGGATGCTTCGCGCTCCCTCTGG	75813
QY	1303	TACACAGGACAGGGCGGATGGG-----CCGCGCGTCTCAGGCGGACC	1347
Db	75814	TACCAACAGACGCGCCCACTGTCGCCCTCGACCTGCACCCGCGGCTGGCCATGCGCGCC	75873
QY	1348	CTGGAGCGGGGAGCAGATCGAGCTAGAGTTCGGTGCCTTTGCGCCAGTTCGCGCATCCTG	1407
Db	75874	CTGCGCGGCCCACTCCGCTCGTACAGCTCCGTGCGCCCTGGTCCAGTTGCGGATGGGC	75933
QY	1408	GTGCGCGGACGCGGCGAGGGGACGAGGCCGATGACAGGAA--CTGCGCGGACGAGG	1464
Db	75934	GTGCTCTCTTGGCGCAGGGGACGAGGCCGATGACGTGAGGGCGGATGCGCGGAAG	75993
QY	1465	AGGGCGTTAGCGAGGACATCCCGGGACCGACGCGCAGGTGGGTTCGATCGATCGCC	1524
Db	75994	AGGGCGTGGGCGAAGGACATCCCGGGCCACGCGGACGCTGTGAGCTCGATCGCG	76053
QY	1525	TGTGAGCGGCTCTGTGCGGCTCTCCCATCGAGCTGCGGGGAGAGCGCGAGAGTGGAG	1584
Db	76054	CGGTGGAGGGGCTCCGCGCTCTTCCCAACGGAGTCAAGGAGCAACCGAGGATGCGA	76113
QY	1585	GGGTTGGGGCGCATCCCGGGGGAACATCCGCTCCACATGGGAGCCGACACGCCCCCGC	1644
Db	76114	GGTGACGGGGCGACTCCCGGGGACACGCGCTCCACGGCCCGTGGTGGCGCCCGC	76173
QY	1645	CGCCCCGCATGTTCGACTGCGCTCCCAAGGATGACACCACTTTTCGACGAGGAGG	1704
Db	76174	CGACCGCCCATGTGTGACTGCGCGCCGCCAAAAGGAATATCAGCGTCCGCGTCGCGC	76233
QY	1705	GACACGCCCCCGCCCGCGGAGGACGACAGCAGCAGCAGCATCCGCGCA	1759
Db	76234	GTGGCGGTGGCCCTGTGTGCGAGCAGCAGCAGCGCTAGTACCTGCAGCTGTA	76288

## RESULT 10

QY	2514	AGTACATGTAACAACACTCGAGACACTCATGAACTCCGAGATGTCCTCACACCGAGAAC	2573
Db	77146	AGTAC-----CACAACTCAAGACCCGATAAACTCCGAGGTGACCTCAGCTGAGAAAT	77200
QY	2574	GACCGAC-----TTGCTTTTATCCTTGGTGGCCCTTGACCTCATCTCCCTCAGCTCGC	2626
Db	77201	GAACCAATATGCAATACCATNTGTCGTTGTGGCCCTGACCTCATCTCCCTTCCCTCGAC	77260
QY	2627	GTCTCTCCTCTTGTCCCTCCCGACGACAAAATGGTGAOCTGCTGTGCGCAGTTAAGT	2686
Db	77261	CTCCTCGCTCTTCTCTCCCTCCCGACGCAATGCGCAACTCAACACCGCAGTTAAGT	77320
QY	2687	TTCTACTTCCGCAATTCACGATTCACACAGCTACATAACCTCTGTTAGATGTCACATGCC	2746
Db	77321	TCTCTCTGCGCACTTGAATGATTCATCAACATTAACCTCTGTTAGATGTCACATGCC	77380
QY	2747	CATTGATGTGGTGTGGGGAAGGGAAGAGATGTGGATGTATCATC-GTCTTGGGGATGA	2805
Db	77381	CAATAAATGCGGTAGGGAAGCGGAGAGAAATGGACATGCCATCAGTTTAGGATGA	77440
QY	2806	TGGCGATCGTCTCTCAGCAAAAGGAGAGTGGGATGTAATGAGATGGAAGGAGGATGATT	2865
Db	77441	TGGCGGTGCTCTCTGCGAAGGAGGTTGGATTGATTGAGTGGAGG-----	77489
QY	2866	GGAGTGAAGGAGTTGGGAGGACCAAGTGCACGACGAGAGAAATGATCTCGAGAGGAGAGA	2925
Db	77490	-----AGTTGGGGAGGAGGACAGGTGGAGATGGAGAAATGATCTCGAGAGGAGAGA	77542
QY	2926	RAGAGGAGAAACAGATATGGAGATAA--GAAATAAATCTAGGCTGTTCTTTAAAGTT	2983
Db	77543	AGGAGAAATAGTTCTAGGGAAGAGAAACGGAAGGATGTGGGTGTTCAATTATTAAGTA	77602
QY	2984	GTATCCATCTCTACGGGAT 3002	
Db	77603	GTATCTATCTCATGGAT 77621	
RESULT	9		
LOCUS	AC096690	145999 bp	DNA linear HTG 08-MAY-2002
DEFINITION	Oryza sativa chromosome 3 clone OSUNBa0059G06, *** SEQUENCING IN		
ACCESSION	AC096690	PROGRESS ***, 5 ordered pieces.	
VERSION	AC096690.4	GI:20503077	
KEYWORDS	HTG: HTGS_PHASE2.		
SOURCE	Oryza sativa.		
ORGANISM	Oryza sativa		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzoideae; Oryza.		
AUTHORS	1 (bases 1 to 145999)		
TITLE	Buell, C.R., Yuan, Q., Ouyang, S., Liu, J., Tallon, L., Gansberger, K., Kim, M., Overton, I.R., Bera, J., Isitrin, T., Krol, M., Jarrah, B., Jin, S., Koo, H., Zismann, V., Hsiao, J., Blunt, S., Vanaken, S., Uterbach, R., Feldblum, T., Yang, Q., Haas, B., Suh, B., Peterson, J., Quackenbush, J., White, O., Salzberg, S. and Fraser, C.		
TITLE	Oryza sativa ssp. japonica cv. Nipponbare OSUNBa0059G06 BAC genomic sequence		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 145999)		
AUTHORS	Buell, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (22-SEP-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA		
REFERENCE	3 (bases 1 to 145999)		
AUTHORS	Buell, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (08-MAY-2002) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA		
COMMENT	On May 8, 2002 this sequence version replaced gi:19881803.		
	* NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. Gaps between the contigs		



```

AP004130
LOCUS      121752 bp      DNA      linear      HTG 21-MAR-2002
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 8 clone
            OJ1014_B05, *** SEQUENCING IN PROGRESS ***, in ordered pieces.
ACCESSION  AP004130
VERSION     HTG; HTGS_PHASE2.
KEYWORDS    Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
SOURCE      clone:OJ1014_B05.
ORGANISM    Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
AUTHORS     Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE       Oryza sativa nipponbare(GA3) genomic DNA, chromosome 8, BAC
            clone:OJ1014_B05
JOURNAL     Published Only in Database (2001)
REFERENCE   2 (bases 1 to 121752)
AUTHORS     Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE       Direct Submission
JOURNAL     Submitted (05-SEP-2001) Takuji Sasaki, National Institute of
            Agrobiological Sciences, Rice Genome Research Program; Kannondai
            2-1-2, Tsukuba, Ibaraki 305-8602, Japan
            (E-mail:tsasakienias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
            Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT     The nucleotide sequence of this BAC clone was generated by
            combining Monsanto and RGP-Japan sequencing data.
            NOTE: It currently consists of 1 contigs. Gaps between the contigs
            are represented as runs of N. The order of the pieces is believed
            to be correct as given, however the sizes of the gaps between them
            are based on estimates that have provided by the submitter. This
            sequence will be replaced by the finished sequence as soon as it is
            available and the accession number will be preserved.
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
FEATURES     source
            1..121752
               /organism="Oryza sativa (japonica cultivar-group)"
               /cultivar="Nipponbare"
               /db_xref="taxon:39947"
               /chromosome="8"
               /clone="OJ1014_B05"
BASE COUNT   32506 a 27609 c 27315 g 34122 t 200 others
ORIGIN
Query Match      4.6%; Score 231; DB 2; Length 121752;
Best Local Similarity 94.1%; Pred. No. 2.5e-28;
Matches 240; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
Qy 4131 TAGTACTCCCTCCGCTCAAAAATAAGTGTAGTTTGTAGCATTATTCATGTTCAACGTTTGA 4190
            |||||||||||
Db 95625 TACTACTCCCTCCGCTCCAAAATAAGTGTAGTTTGTAGCATTATTCATGTTCAACGTTTGA 95684
Qy 4191 CCATTGCGTCTATTGTGAAAAGAGATATGATTAGTATTTTATTGTTATTAGATGATAAA 4250
            |||||||||||
Db 95685 CCGTTGCGTCTATTGTGAAAATAATATTATGATTAGTATTTTGTGTTATTAGATGATAAA 95744
Qy 4251 ACATGAATAGTACTTTATGTCGACTAAATTTTATTTTATTTTATTTTATTTTCAAA 4310
            |||||||||||
Db 95745 ACATGAATAGTAAATTATGTCGACTAAATTTTATTTTATTTTATTTTCAAA 95804
Qy 4311 ATAAGACGGATGGTCAAAAGCGCTAAACATGGATATCTATGCGTACACTTTATTTGGGAGC 4370
            |||||||||||
Db 95805 ATAAGACGGATGGTCAAAAGCGCTAAACATGGATATCTATGCGTACACTTTATTTGGGAGC 95864
Qy 4371 GAGGTAGTAGGCTAT 4385
            |||||||||
Db 95865 GAGGTAGTAGATTTT 95879

```

```

RESULT 11
LOCUS      176224 bp      DNA      linear      HTG 17-MAY-2002
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 8 clone
            OSJNBa0053M06, *** SEQUENCING IN PROGRESS ***, in ordered pieces.
ACCESSION  AP005163
VERSION     GI:20975289
KEYWORDS    HTG; HTGS_PHASE2.
SOURCE      Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
            clone:OSJNBa0053M06.
ORGANISM    Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
AUTHORS     Sasaki,T., Matsumoto,T. and Katayose,Y.
TITLE       Oryza sativa nipponbare(GA3) genomic DNA, chromosome 8, BAC
            clone:OSJNBa0053M06
JOURNAL     Published Only in Database (2002)
REFERENCE   2 (bases 1 to 176224)
AUTHORS     Sasaki,T., Matsumoto,T. and Katayose,Y.
TITLE       Direct Submission
JOURNAL     Submitted (15-MAY-2002) Takuji Sasaki, National Institute of
            Agrobiological Sciences, Rice Genome Research Program; Kannondai
            2-1-2, Tsukuba, Ibaraki 305-8602, Japan
            (E-mail:tsasakienias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
            Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT     The nucleotide sequence of this BAC clone consists of 1 contigs. Gaps between the contigs
            are represented as runs of N. The order of the pieces is believed
            to be correct as given, however the sizes of the gaps between them
            are based on estimates that have provided by the submitter. This
            sequence will be replaced by the finished sequence as soon as it is
            available and the accession number will be preserved.
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
FEATURES     Location/Qualifiers
            1..176224
               /organism="Oryza sativa (japonica cultivar-group)"
               /cultivar="Nipponbare"
               /db_xref="taxon:39947"
               /chromosome="8"
               /clone="OSJNBa0053M06"
BASE COUNT   49574 a 38250 c 38489 g 49729 t 182 others
ORIGIN
Query Match      4.6%; Score 231; DB 2; Length 176224;
Best Local Similarity 94.1%; Pred. No. 2.5e-28;
Matches 240; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
Qy 4131 TAGTACTCCCTCCGCTCAAAAATAAGTGTAGTTTGTAGCATTATTCATGTTCAACGTTTGA 4190
            |||||||||||
Db 46774 TACTACTCCCTCCGCTCCAAAATAAGTGTAGTTTGTAGCATTATTCATGTTCAACGTTTGA 46833
Qy 4191 CCATTGCGTCTATTGTGAAAAGAGATATGATTAGTATTTTATTGTTATTAGATGATAAA 4250
            |||||||||||
Db 46834 CCGTTGCGTCTATTGTGAAAATAATATTATGATTAGTATTTTGTGTTATTAGATGATAAA 46893
Qy 4251 ACATGAATAGTACTTTATGTCGACTAAATTTTATTTTATTTTATTTTCAAA 4310
            |||||||||||
Db 46894 ACATGAATAGTAAATTATGTCGACTAAATTTTATTTTATTTTATTTTCAAA 46953
Qy 4311 ATAAGACGGATGGTCAAAAGCGCTAAACATGGATATCTATGCGTACACTTTATTTGGGAGC 4370
            |||||||||||
Db 46954 ATAAGACGGATGGTCAAAAGCGCTAAACATGGATATCTATGCGTACACTTTATTTGGGAGC 47013
Qy 4371 GAGGTAGTAGGCTAT 4385
            |||||||||
Db 47014 GAGGTAGTAGATTTT 47028

```

RESULT 12

AP004224 152835 bp DNA linear PLN 13-JUL-2002  
 LOCUS Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,  
 DEFINITION BAC clone:B1111E11.  
 ACCSSION AP004224.3 GI:21743333  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,  
 clone:B1111E11.  
 Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 1  
 Sasaki,T., Matsumoto,T. and Yamamoto,K.  
 Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC  
 clone:B1111E11  
 Published Only in Database (2001)  
 2 (bases 1 to 152835)  
 Sasaki,T., Matsumoto,T. and Yamamoto,K.  
 Direct Submission  
 Submitted (03-OCT-2001) Takuji Sasaki, National Institute of  
 Agrobiological Sciences, Rice Genome Research Program, Kannondai  
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
 (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,  
 Tel:81-298-38-7441, Fax:81-298-38-7458)  
 On Jul 12, 2002 this sequence version replaced gi:17026098.  
 Genes were predicted from the integrated results of the following:  
 GENSCAN1.0, ELASTN2.0, ELASTX2.0 as well as SplicePredictor  
 (October 1998 version). The genomic sequence was searched against  
 NCBI NonRedundant Protein database, nr  
 (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at  
 RGP. Protein homologs of the coding regions were searched against  
 NCBI NonRedundant Protein database with BLASTp2.0. ESTs represent  
 the identified cDNA sequences using ELASTN 2.0 with the  
 corresponding DBJ accession no. and RGP clone ID.  
 A gene with identity or significant homology to a protein is  
 classified based on the protein name to indicate the homology level  
 such as same name, 'putative-' and '-like protein'. A gene without  
 significant homology to any protein but with EST homology (covering  
 almost the entire length of partial sequence) is classified as an  
 'unknown' protein. A gene predicted with a gene prediction program  
 is classified as a 'hypothetical' protein.  
 The orientation of the sequence is from -21M13 to M13-rev of the BAC  
 clone. This sequence of B1111E11 clone has an overlap with B1080D07  
 (DD87:AP003203) at the position 146,696 to 152,835 of 3' end.  
 Detailed information on overlap and assembly quality together with  
 annotation of this entry is available at  
 http://rgp.dna.affrc.go.jp/GenomeSeq.html.  
 FEATURES  
 Location/Qualifiers  
 1. 152835  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /chromosome="1"  
 /clone="B1111E11"  
 complement(1..514)  
 /note="3' LTR"  
 /gene="B1111E11.1"  
 complement(530..6029)  
 /gene="B1111E11.1"  
 complement(530..6029)  
 /gene="B1111E11.1"  
 /note="probably inactive due to stop codon(s) in CDS  
 pseudogene, polyprotein"  
 /pseudo  
 Join(8253..8336,8386..8595)  
 /gene="B1111E11.2"  
 Join(8253..8336,8386..8595)  
 /gene="B1111E11.2"  
 /note="hypothetical protein"  
 /codon\_start=1  
 /protein\_id="BAC03327.1"  
 /db\_xref="GI:21743334"

/translation="MTSPPAVARARKLASERRRDDANGGHOHNGRRRRGRNGNLR  
 LDDGGAPANYGGGVDEIGDGLVTTATSPNSDDRTDGGARLKRRRRR"  
 complement(8839..9352)  
 /note="5' LTR"  
 complement(Join(14589..14792,14904..15023,15078..15290,  
 15389..15607,15720..15865,16296..16365))  
 /gene="B1111E11.3"  
 complement(Join(14589..14792,14904..15023,15078..15290,  
 15389..15607,15720..15865,16296..16365))  
 /gene="B1111E11.3"  
 /codon\_start=1  
 /product="putative annexin"  
 /protein\_id="BAC03328.1"  
 /db\_xref="GI:21743335"  
 /translation="MEDHTRKLNQKADIASAEVHKSGGWGDEQAVIGVLAHRDATOR  
 KQIRLTENYNENLQRLQSLGDLERAHWLPDVERQAVVNTATKCIHEDYA  
 VVVIACNTSSSELAVKETYHYVQLQSLSEDAARATNLRLSLALVYRSDGE  
 VNDALAKSEAKILHETVNGTDHGELIRIVGTRSAQLNATFSWFRDERGTSIKIH  
 APRFDQALQGDADPTGYSHALRTALCISDANKYFVKLRNMAHKSGTNEDSLTRVI  
 VLHAKDLGIGKDAFOKRAVALEKAIGNDTSGDYKSELMALIGSGI"  
 complement(Join(17303..17441,21169..21190,21797..22025,  
 22151..22324))  
 /gene="B1111E11.4"  
 complement(Join(17303..17441,21169..21190,21797..22025,  
 22151..22324))  
 /gene="B1111E11.4"  
 /note="hypothetical protein"  
 /codon\_start=1  
 /protein\_id="BAC03329.1"  
 /db\_xref="GI:21743336"  
 /translation="MVYAMIDSIRLINKYQKRDQVORNFRLDITMIDILKIVIRM  
 KNTVYVSLILGVGILLINSYSGEICVCKYKETHPAYKETHLRPTRKYTQWRA  
 HKSGDTRERWIVITAIRANOSVSLPKERGLDSEVDGLSKNPKMLKINIALDQ  
 WINVHIFTRINITEINAGKRIPIVSG"  
 Join(24567..24773,24885..25163,25427..25524,25675..25882)  
 /gene="B1111E11.5"  
 Join(24567..24773,24885..25163,25427..25524,25675..25882)  
 /gene="B1111E11.5"  
 /note="hypothetical protein  
 similar to Oryza sativa chromosome10, OSJNB00089F16.9"  
 /codon\_start=1  
 /protein\_id="BAC03330.1"  
 /db\_xref="GI:21743337"  
 /translation="MTGFNLSENFKENPEAFRRVPRVVPVAPKSLPTKLAIPVPT  
 EKTWASLTLEFAAPSADNVATLOFLIQCSTYTIKGNLDAYRLRLEPLFSLIGRAQ  
 WFTYNVAVTWKCSATFLRSFFPKINALRGRISSFOQTRDESIEPNERLQETG  
 KWTLEBRAGGTRETRNRRNRESSTSPVOMSSGPPSPSESSSSSSNDLSEI  
 IRRMDTLDIQIGEIQYNLTETHVAQTEWQQAQAFANINDMMQO"  
 30135..32223  
 /gene="B1111E11.6"  
 30135..32223  
 /gene="B1111E11.6"  
 /note="probably inactive due to stop codon(s) in CDS  
 pseudogene, retroelement"  
 /pseudo  
 Join(35262..35643,43596..43702)  
 /gene="B1111E11.7"  
 Join(35262..35643,43596..43702)  
 /gene="B1111E11.7"  
 /note="hypothetical protein"  
 /codon\_start=1  
 /protein\_id="BAC03331.1"  
 /db\_xref="GI:21743338"  
 /translation="MFLTGKVKRQRMIFGHIGVGVCRGKNSLRSHGELPSINA  
 NIGTRGCDVFRNQAOKMPKDFWTYWSLRAFAVTHVFAANFOLLATLAHGG  
 AMFFDRNEKVQKAPKQDFWTYWSVKERTAFETFGDDREPFENLRNDRSAKHTSKAT  
 SL"  
 Join(44530..44843,44934..45375)  
 /gene="B1111E11.8"  
 Join(44530..44843,44934..45375)  
 /gene="B1111E11.8"  
 /note="hypothetical protein  
 similar to Oryza sativa chromosome3, OSJNBa0026A15.9"

```

/codon_start=1
/protein_id="BAC03332.1"
/db_xref="GI:21743339"
/translation="MRRRSQQLHGGECAHAAANPTTAAAPRDPDEPARRNGREE
GMATSDGDDHRRSGSGGEARPRATPLKCRWRWRGRRMGSGGGEARNGPGETGE
WPGSGDATDRHEHAGSLRAKRRRRWRDCEAEEGGVDVSHRTGEAAQAAGGWP
PREKEVEGRGAATENLGKLEKRENRARGFHEIGERELATGEANGTGGGIGARRWK
VAGVGAGVDDMGHESRGKLGTEEGXGINSF"
complement(46012..46653)
/gene="B1111E11.9"
CDS
complement(46012..46653)
/gene="B1111E11.9"
/notes="hypothetical protein
similar to Oryza sativa chromosome10, OSJNB0089F16.9"
/codon_start=1
/protein_id="BAC03333.1"
/db_xref="GI:21743340"
/translation="MTGFNLPENFKYEPFAFRGVRPRAPVAPQKLTPEKPAIPAPI
FTWASKTLREFAEDTNVAIGPQINIGDMDFDLKSLITMAQASFCKGPNVDANA
HLQFLEICSTYTIKGVSPDAIRLRFPSLIGRAKQFYANRAVNTWKSPGLS
KFPFMKTWALGRISNFQOTRDSIPEALERLQEVAACPHGHGMDWLIVTS"
complement(48388..48579)
/gene="B1111E11.10"
CDS
complement(48388..48579)
/gene="B1111E11.10"
/notes="hypothetical protein"
/codon_start=1
/protein_id="BAC03334.1"
/db_xref="GI:21743341"
/translation="MRALPEKYELVTMLNSDMRMTFASLLGKINTNDMYKLKK
EMEAASFAPKUKIRAK"
53847..54335
/notes="5' LTR"
55770..57616
/gene="B1111E11.11"
55770..57616
/gene="B1111E11.11"
/notes="B1111E11.11"
/notes="probably inactive due to stop codon(s) in CDS
pseudogene, retroelement"
/pseudo
57629..58116
/notes="3' LTR"
join(60720..61097,61272..61421,62316..62685,62854..62971,
63077..63192,63282..63374,63454..63549,64125..64847,
64968..65116,66390..66476,66497..66581)
/gene="B1111E11.12"
join(60720..61097,61272..61421,62316..62685,62854..62971,
63077..63192,63282..63374,63454..63549,64125..64847,
64968..65116,66390..66476,66497..66581)
/gene="B1111E11.12"
/notes="hypothetical protein
similar to Arabidopsis thaliana chromosome 3, At3g54500"
/codon_start=1
/protein_id="BAC03335.1"
/db_xref="GI:21743342"

```

```

Query Match 4.4%; Score 217.8; DB 8; Length 152835;
Best Local Similarity 86.6%; Pred. No. 4e-26;
Matches 252; Conservative 0; Mismatches 37; Indels 2; Gaps 1;

QY 4129 AATAGTACTCCCTCGCTCCAAATAAGTGTAGTATTAGCACTATTTCATGTTCAACGTTT 4188
DB 78155 AAGACTACTCCCTCGCTCCCAANAATAAGTGTAGTATTTCATGTTCAACGTTT 78214
QY 4189 GACCATTCGCTTATTGAAAAGAATATGATTAGTATTTTATTGTTATTAGATGATA 4248
DB 78215 GACCGTTCATCTATTGAAAAGAATATGATTAGTATTTTATTGTTATTAGATGATA 78274
QY 4249 AACATGAAATAGTACTTATGCTAGCTAAATTTTTTTTAAATATTTTATTAATTTTC 4308
DB 78275 AACATGAAATAGTACTTATGCTAGCAAA--TTTTTTAAATATTTTTCATAAATTTTC 78332
QY 4309 AATAAGCGGATGTCACAAAGCGCTAAACATGATATCTATGGCTACACATTTTGGGA 4368

```

```

|||||
Db 78333 AATAGACGATGGCTCAAGCGCTAAACAGGATATCTATGCTCCACTTATTTGGGA 78392
|||||
QY 4369 CGGAGTAGTAGGCTATATAACACAGCTATATAACATATTTTAAAGAGATAAAA 4419
|||||
Db 78393 CGGAGTAGTATACCTAGTATCCCAACCAATAAGATCTAACTTATCGAA 78443
|||||

RESULT 13
AP004849/c 163704 bp DNA linear HTG 23-MAR-2002
LOCUS Oryza sativa (japonica cultivar-group) chromosome 2 clone
DEFINITION OJ1311.D08, ** SEQUENCING IN PROGRESS ***, in ordered pieces.
ACCESSION AP004849
VERSION AP004849.1 GI:19698269
KEYWORDS HTG; HTGS_PHASE2.
SOURCE Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
clone:OJ1311.D08.
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1
Sasaki,T., Matsumoto,T. and Yamamoto,K.
Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, BAC
clone:OJ1311.D08
JOURNAL Published Only in Database (2002)
REFERENCE 2 (bases 1 to 163704)
Sasaki,T., Matsumoto,T. and Yamamoto,K.
Direct Submission
Submitted (20-MAR-2002) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
The nucleotide sequence of this BAC clone was generated by
combining Monsanto and RGP-Japan sequencing data.
NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: this is a 'working draft' sequence.
* this sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
Location/Qualifiers
source 1..163704
/organism="Oryza sativa (japonica cultivar-group)"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosomes="2"
/clone="OJ1311.D08"
BASE COUNT 46733 a 35204 c 34892 g 46875 t
ORIGIN
Query Match 4.3%; Score 216.6; DB 2; Length 163704;
Best Local Similarity 92.0%; Pred. No. 6.3e-26;
Matches 240; Conservative 0; Mismatches 19; Indels 2; Gaps 1;

QY 4129 AATAGTACTCCCTCGCTCCAAATAAGTGTAGTATTAGCACTATTTCATGTTCAACGTTT 4188
DB 77913 ACTGTACTCCCTCGCTCCCAANAATAAGTGTAGTATTAGCAATATTCATGTTTACATTT 77954
QY 4189 GACCATTCGCTTATTGAAAAGAATATGATTAGTATTTTATTGTTATTAGATGATA 4248
DB 77953 GACCGTTCGCTTATTGAAAAGAATATGATTAGTATTTTATTGTTATTAGATGATA 77794
QY 4249 AACATGAAATAGTACTTATGCTAGCTAAATTTTTTTTAAATATTTTATTAATTTTC 4308
DB 77793 AACATGAAATAGTACTTATGCTAGCTAA--TTTTTTAAATATTTTATTAATTTTC 77736

```

```

QY 4309 AATAAGACGAGTGTCRAAGCGCTAAACATGGATATCTATGCTACACTATTATTGGGA 4368
|||||
DB 77735 AATAGACGAGTGCCTAAGTGCTTAATATGGAATATGCTATGCTACACTATTATTGGGA 77676
|||||

QY 4369 CGGAGGTAGTATGCTTCCAAAC 4389
|||||
DB 77675 CGGAGGTAGTATGCTTCCAAAC 77655
|||||

RESULT 14
AC118980
LOCUS
DEFINITION
Oryza sativa chromosome 3 clone OJ1263H11, linear HTG 23-APR-2002
PROGRESS ***, 3 ordered pieces.
ACCESSION
AC118980
VERSION
AC118980.1 GI:20270052
KEYWORDS
HTG; HTGS_PHASE2.
SOURCE
Oryza sativa.
ORGANISM
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
Wing,R.A., Yu,Y., Soderlund,C., Chen,M., Kim,H.-R., Rambo,T.,
Saski,C., Henry,D., Oates,R. and Simmons,J.
1 (bases 1 to 126637)
TITLE
Rice Genomic Sequence
JOURNAL
Unpublished
REFERENCE
2. (bases 1 to 126637)
Wing,R.A., Yu,Y., Soderlund,C., Chen,M., Kim,H.-R., Rambo,T.,
Saski,C., Henry,D., Oates,R. and Simmons,J.
Direct Submission
TITLE
Submitted (23-APR-2002) Clemson University Genomics Institute,
Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA
JOURNAL
COMMENT
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
1 19607: contig of 19607 bp in length
*
* 19608 90592: contig of 70985 bp in length
*
* 90593 126637: contig of 36045 bp in length.
*
Location/Qualifiers
1. .126637
/organism="Oryza sativa"
/db_xref="taxon:4530"
/chromosome="3"
/clone="OJ1263H11"
BASE COUNT 36868 a 27251 c 27250 g 35263 t 5 others
ORIGIN
Query Match 4.3%; Score 215.8; DB 2; Length 126637;
Best Local Similarity 94.4%; Pred. No. 8.4e-26;
Matches 235; Conservative 0; Mismatches 12; Indels 2; Gaps 1;

QY 4131 TAGTACTCCCTCCGCTCCAAATAAGTGTAGTTTACGACTATCATGTTCAAGCTTTGA 4190
|||||
DB 113552 TAGTACTCCCTCCGCTCCCAATAAGTGCAGTTTACGACTATTCATGTTCAAGCTTTGA 113611
|||||

QY 4191 CCATTCGCTCTATTGAAAAGAGATATGATTAGTATTATTGTTATAGATGATAAA 4250
|||||
DB 113612 CGGTCTGCTCTATTGAAAAGAGATATGATTAGTATTATTGTTATAGATGATAAA 113671
|||||

QY 4251 ACATGATAGTACTTATGTGTGCACTAATTTTTTTTAAATATTTTATTAATTTTCAA 4310
|||||
DB 113672 ATAAGAATAGTACTTATGTGTGAATACTAA--TTTTTTTAAATATTTTATTAATTTTCAA 113729
|||||

QY 4311 ATAAGAAGGATGTCRAAGCGCTAAACATGGATATCTATGCTACACTATTATTGGGACG 4370
|||||

```

```

|||||
DB 113730 ATAAGACGAGTGCCTAAGCGCTAAATACGATATTTATGGCTGCACACTATTATTGGGACG 113789
|||||

QY 4371 GAGTAGTA 4379
|||||
DB 113790 GAGTAGTA 113798
|||||

RESULT 15
AP003946
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) chromosome 6 clone
OJ1147.D11, *** SEQUENCING IN PROGRESS ***, in ordered pieces.
ACCESSION
AP003946
VERSION
AP003946.1 GI:15021916
KEYWORDS
HTG; HTGS_PHASE2.
SOURCE
Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
clone:OJ1147.D11.
ORGANISM
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
Sasaki,T., Matsumoto,T. and Yamamoto,K.
1
Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, BAC
clone:OJ1147.D11
Published Only in Database (2001)
2 (bases 1 to 138906)
Sasaki,T., Matsumoto,T. and Yamamoto,K.
Direct Submission
TITLE
Submitted (25-JUL-2001) Takuji Sasaki, National Institute of
Agricultural Resources, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasakienias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
The nucleotide sequence of this BAC clone was generated by
combining Monsanto and RGP-Japan sequencing data.
NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces are believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
Location/Qualifiers
1. .138906
/organism="Oryza sativa (japonica cultivar-group)"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="6"
/clone="OJ1147.D11"
BASE COUNT 39967 a 29996 c 29997 g 38791 t 155 others
ORIGIN
Query Match 4.3%; Score 215.4; DB 2; Length 138906;
Best Local Similarity 91.3%; Pred. No. 9.9e-26;
Matches 240; Conservative 0; Mismatches 21; Indels 2; Gaps 1;

QY 4117 AAGAGCAGGTAGTAATAGTACTCCCTCCGCTCCAAATAAGTGTAGTTTACGACTATTCA 4176
|||||
DB 11539 AGGGGAAGAAACAAGTACTCCCTCCGCTCTTAATAAGTGCAGTTTACGACTGTCA 11598
|||||

QY 4177 TGTTCAACGTTTGACCATTCGCTTTATTGTTGAAAAGAGATATGATTAGTATTTTATGT 4236
|||||
DB 11599 TGTTCAACGTTTGACCGTCTCTATTGTTGAAAAGATATGATTAGTATTTTATGT 11658
|||||

QY 4237 TATTAGATGATAAAACATGAATAGTACTTTATGCTGCTACTATTTTAAATATTTT 4296
|||||
DB 11659 TATTAGATGATAAAACATGAATAGTACTTTATGCTGCTACTATTTTAAATATTTT 11716
|||||

```

## FEATURES

source



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2003, 18:27:39 ; Search time 762.288 Seconds  
(without alignments)  
14774.266 Million cell updates/sec

Title: US-09-702-134-7212\_COPY\_62300\_67300

Sequence: 1 aggggacactaacttoca.....tgccgcggagtgccgcccg 5001

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002.\*

1:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*	
2:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*	
3:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*	
4:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*	
5:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*	
6:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*	
7:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*	
8:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*	
9:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*	
10:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*	
11:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*	
12:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*	
13:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*	
14:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*	
15:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*	
16:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*	
17:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*	
18:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*	
19:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*	
20:	/SID32/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*	
21:	/SID32/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*	
22:	/SID32/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*	
23:	/SID32/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*	
24:	/SID32/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	139.8	2.8	561	24 ABL49988	Rice Ditto-Os1 nuc
2	139.8	2.8	561	24 ABL49989	Rice Ditto-Os1 nuc
3	137.4	2.7	5113	21 AA261434	Nucleotide sequenc
4	129.8	2.6	544	24 ABL49987	Rice Ditto-Os1 nuc
5	121.8	2.4	5757	21 AA261433	Nucleotide sequenc
c 6	118.8	2.4	2943	21 AAC48061	Arabidopsis thalia
7	118.4	2.4	949	24 ABN98574	Arabidopsis thalia
c 8	118.4	2.4	1024	21 AAC45528	Arabidopsis thalia
c 9	118.4	2.4	3108	21 AAC46278	Arabidopsis thalia

c 10	118.4	2.4	3187	21 AAC47620	Arabidopsis thalia
c 11	117.8	2.4	1024	21 AAC4895	Arabidopsis thalia
c 12	114.4	2.3	547	21 AA79573	Pinus radiata cell
c 13	107.8	2.2	2380	20 AAX32000	Rice pyruvate deca
c 14	107.8	2.2	5526	20 AAX31999	Rice pyruvate deca
c 15	99	2.0	15686	24 AAD38801	Rice RGA38 contig.
c 16	94.4	1.9	5579	24 AAL46958	Rice lesion inhibi
c 17	91.6	1.8	1730	24 AAI72980	OSGRP-Al full leng
c 18	91.6	1.8	3572	24 AAI72979	OSGRP-Al coding se
c 19	89.6	1.8	267	24 ABL49990	Rice Ditto-Os2 nuc
c 20	84.2	1.7	312	24 ABL49992	Rice Ditto-Os2 nuc
c 21	83.4	1.7	246	24 ABL49991	Rice Ditto-Os2 nuc
c 22	75.6	1.5	500	24 AAI72978	OSGRP-Al larger pr
c 23	73.2	1.5	1026	20 AAX3542	Rice beta-glucanas
c 24	73.2	1.5	2169	20 AAX3534	Rice beta-glucanas
c 25	69.6	1.4	91552	24 AAD38803	BAC clone KGP36 fr
c 26	65.4	1.3	18155	24 AAD38808	CODR4 ORF from ric
c 27	65.4	1.3	91552	24 AAD38803	BAC clone KGP36 fr
c 28	64.6	1.3	4569	24 ABK15663	Rice lipoxigenase
c 29	62.8	1.3	1993	21 AA63731	DNA encoding a pol
c 30	62.4	1.2	58857	21 AA58471	Nucleotide sequenc
c 31	62	1.2	1806	21 AA50102	Human DMF1 DNA. Or
c 32	62	1.2	12120	24 AAS96695	Rice DMF1 DNA. Or
c 33	62	1.2	125401	22 AAD17186	Streptomyces nous
c 34	61.6	1.2	65140	22 AAD17184	Streptomyces nous
c 35	61.2	1.2	1050	22 AAD21685	Mutational hot spo
c 36	61.2	1.2	2849	22 AAD21684	Human retinitis pi
c 37	61	1.2	863	24 ABK77958	Bacillus clausii g
c 38	61	1.2	4403765	22 AAI99583	Mycobacterium tube
c 39	61	1.2	4411529	22 AAI99582	Mycobacterium tube
c 40	60.6	1.2	1280	21 ABQ62495	Mycobacterium bovi
c 41	60.6	1.2	15872	18 AAT68715	Streptomyces venez
c 42	60.6	1.2	15872	21 AA287283	S. venezuelae vep
c 43	60.4	1.2	1591	18 AAT93780	Oryza sativa patho
c 44	60	1.2	390	13 AAQ21833	Randomising oligon
c 45	60	1.2	390	14 AAQ36859	PCR primer for 5'

## ALIGNMENTS

RESULT 1	
ABL49988	
ID	ABL49988 standard; DNA; 561 BP.
AC	ABL49988;
XX	
DT	10-JUN-2002 (first entry)
XX	
DE	Rice Ditto-Os1 nucleotide sequence SEQ ID NO:63.
XX	
KW	Rice; plant; detection; polymorphism; transposable element; genome;
KW	gene; ds.
XX	
OS	Oryza sativa.
XX	
PN	WO200212484-A1.
XX	
PD	14-FEB-2002.
XX	
PF	02-AUG-2001; 2001WO-JP066661.
XX	
PR	02-AUG-2000; 2000JP-0234577.
XX	
PA	(NLSB ) JAPAN TOBACCO INC.
PA	(SYGN ) SYNGENTA LTD.
XX	
PI	Komori T, Nitta N;
XX	
DR	WPI; 2002-241759/29.
XX	
PT	Construction of marker for detecting plant genome polymorphism with use of transposable element, useful particularly in studying restriction



```

FT      /*tag= b
FT      /number= 1
FT      intron 1592..3056
FT      /*tag= c
FT      exon   3057..4898
FT      /*tag= d
FT      /number= 2
XX      WO200008162-A1.
XX      17-FEB-2000.
XX      03-AUG-1999; 99WO-US17706.
XX      04-AUG-1998; 98US-0095229.
XX      21-JUN-1999; 99US-0336946.
XX      (DUPO ) DU PONT DE NEMOURS & CO E I.
XX      PA
XX      PI Valent BS, Bryan GT;
XX      DR WPI; 2000-205715/18.
XX      DR P-PSDB; AAY69308.
XX      PS
XX      Novel nucleic acid fragments conferring Pi-ta resistance gene-mediated
XX      defence response for producing transgenic plants resistant to fungal
XX      pathogens, especially rice blast fungus -
XX      Disclosure; Page 60-62; 96pp; English.
XX      The present sequence encodes a disease resistance protein of rice
XX      variety Tsuyuake. The rice Pi-ta gene was cloned by a map-based
XX      cloning strategy. The pi-ta protein has a novel structure, compared
XX      to all known classes of resistance gene products. The polynucleotide
XX      sequence confers a Pi-ta resistance gene-mediated defence response
XX      against diseases caused by fungal pathogens, particularly the rice
XX      blast fungus. Introduction of the cloned Pi-ta gene into susceptible
XX      rice confers resistance to pathogen strains.
XX      SQ
XX      Sequence 5113 BP; 1500 A; 1033 C; 1075 G; 1505 T; 0 other;
XX
XX      Query Match 2.7%; Score 137.4; DB 21; Length 5113;
XX      Best Local Similarity 85.5%; Pred. No. 3.2e-19;
XX      Matches 165; Conservative 0; Mismatches 26; Indels 2; Gaps 1;
XX
XX      QY 4375 TAGTAGCTATATATTTAGCCAGCTACAGCGGACTCAAGAGCGTAATGTGTATTA 4434
XX      DB 2630 TAGCGGTTATAGCTAGCTATATATTTTAAAGAGATTAAGAGAGAGAGAGAG 2689
XX      QY 4435 CAGCGGACTATATTTAGCCAGCTACAGCGGACTCAAGAGCGTAATGTGTATTA 4494
XX      DB 2690 CAGCAGCTACAGCTAGCTAGCCAGCTGCGAGCGGACTTAAGACGTAATGTGTG--TA 2747
XX      QY 4495 TCACAGTAGGACCGATTATTAAGAGTATAGTAGAGCACTATTTGTATGAATAGCTATTA 4554
XX      DB 2748 TCACAGTAGGACCGATTATTAAGTATAGTAGAGCACTATTTGTATGAATGTGTATTT 2807
XX      QY 4555 CATTAACTATGAT 4567
XX      DB 2808 GGCTCTAGATGAT 2820
XX
XX      RESULT 4
XX      ABL49987
XX      ID ABL49987 standard; DNA; 544 BP.
XX      AC ABL49987;
XX      XX
XX      10-JUN-2002 (first entry)
XX      DE Rice Ditto-Osl nucleotide sequence SEQ ID NO:62.
XX

```

```

KW      Rice; plant; detection; polymorphism; transposable element; genome;
KW      gene; ds.
XX      OS Oryza sativa.
XX      PN WO200212484-A1.
XX      PD 14-FEB-2002.
XX      XX
XX      PF 02-AUG-2001; 2001WO-JP06661.
XX      XX
XX      PR 02-AUG-2000; 2000JP-0234577.
XX      XX
XX      PA (NLSB ) JAPAN TOBACCO INC.
XX      PA (SYGN ) SYNGENTA LTD.
XX      PI Komori T, Nitta N;
XX      XX
XX      DR WPI; 2002-241759/29.
XX      XX
XX      PT Construction of marker for detecting plant genome polymorphism with use
XX      of transposable element, useful particularly in studying restriction
XX      fragment length polymorphism applicable in cell genetics -
XX      PS
XX      Example 4; Fig 17; 110pp; Japanese.
XX      The present invention describes a method for constructing a marker for
XX      detecting polymorphisms in a plant genome. The method comprises the
XX      production of a primer for nucleic acid amplification by using the base
XX      sequences of a transposable element and/or the domain adjacent to it.
XX      Also described is a marker for detection polymorphisms in plant genomes.
XX      The constructed marker can be used for detecting plant genome
XX      polymorphisms, which is useful particularly in studying restriction
XX      fragment length polymorphism applicable in cell genetics e.g. for
XX      analysing and selecting specific breeds of plants. The method is simple,
XX      easy, less time consuming and not so laborious, e.g. in the study of
XX      less frequently occurring polymorphism between various species to enable
XX      identification of the site and isolation of the required gene.
XX      ABL49926 to ABL50036 represent nucleotide sequences used in the
XX      exemplification of the present invention.
XX      SQ
XX      Sequence 544 BP; 179 A; 72 C; 91 G; 202 T; 0 other;
XX
XX      Query Match 2.6%; Score 129.8; DB 24; Length 544;
XX      Best Local Similarity 74.9%; Pred. No. 5.5e-18;
XX      Matches 176; Conservative 0; Mismatches 57; Indels 2; Gaps 1;
XX
XX      QY 4385 TAAACCGCTATAAACAATATTTTAAAGAGATAAAGAGAGAGAGAGAGAGAGAGAGAG 4444
XX      DB 34 TAGTCAGCTATAAACAATATTTTAAAGAGATAAAGAGAGAGAGAGAGAGAGAGAGAG 93
XX      QY 4445 TATATTTAGCAGCTACAGCAGGACTCAAGAGCGTAATGTGTATATGACAAGTAG 4504
XX      DB 94 CAGATCTGTAGTCAGCTCAGCAGCGGACTTCAATACGTAAATGTATG--TATGATAGGTGG 151
XX      QY 4505 GACCAGCTATTAAGAGTATAGTAAGCAACTATTCGTATGAATTAGCTATTAATTAACAT 4564
XX      DB 152 GACCAGCTATTAAGTACAGTAAGCAACTATTCGTATGAATTAGCTATTAATTAACAT 211
XX      QY 4565 GATGATTTGGAGCGCTCTTAACGAGCGGTGATATTTTATTTAGGCCAACATCTCTCA 4619
XX      DB 212 AGATGATTTTAACTTAGTAGTGGGCTACTATTATAAACTTGCTCTTAATTTGTTTA 266
XX
XX      RESULT 5
XX      AAZ61433
XX      ID AAZ61433 standard; DNA; 5757 BP.
XX      AC AAZ61433;
XX      XX
XX      19-JUN-2000 (first entry)
XX      DE Nucleotide sequence of the genomic clone of the Pi-ta gene.

```



XX	Disease resistance protein; rice; variety Yashiro-mochi; Pi-ta gene;
KW	resistance gene; Pi-ta resistance gene-mediated defence response;
KW	fungal pathogen; rice blast fungus; ss.
XX	
OS	Oryza sativa.
XX	
Key	Location/Qualifiers
CDS	1256..5505
FT	/tag= a
FT	/product=
FT	/note= "contains 1 intron"
FT	exon
FT	1256..2200
FT	/tag= b
FT	/number= 1
FT	intron
FT	2201..3663
FT	/tag= c
FT	/number= 1
FT	exon
FT	3664..5505
FT	/tag= d
FT	/number= 2
XX	
PN	WO200008162-A1.
XX	
PD	17-FEB-2000.
XX	
PF	03-AUG-1999; 99WO-US17706.
XX	
PR	04-AUG-1998; 98US-0095229.
PR	21-JUN-1999; 99US-0336946.
XX	
PA	(DUPO ) DU PONT DE NEMOURS & CO E I.
XX	
PI	Valent BS, Bryan GT;
XX	
WI	WPI: 2000-205715/18.
DR	p-PSDS; AAY69307.
XX	
PT	Novel nucleic acid fragments conferring Pi-ta resistance gene-mediated
PT	defence response for producing transgenic plants resistant to fungal
PT	pathogens, especially rice blast fungus -
XX	
PS	Claim 1; Page 56-57; 96pp; English.
XX	
CC	The present sequence encodes a disease resistance protein of rice
CC	variety Yashiro-mochi. The rice Pi-ta gene was cloned by a map-based
CC	cloning strategy. The pi-ta protein has a novel structure, compared
CC	to all known classes of resistance gene products. The polynucleotide
CC	sequence confers a pi-ta resistance gene-mediated defence response
CC	against diseases caused by fungal pathogens, particularly the rice
CC	blast fungus. Introduction of the cloned pi-ta gene into susceptible
CC	rice confers resistance to pathogen strains.
XX	
SQ	Sequence 5757 BP; 1638 A; 1202 C; 1251 G; 1666 T; 0 other;
XX	
Query Match	2.4%; Score 121.8; DB 21; Length 5757;
Best Local Similarity	86.0%; Fred. No. 7.5e-16;
Matches 147; Conservative	0; Mismatches 22; Indels 2; Gaps
QY	4397 AAACATATTTTAAAGAGATAAAGAAGAGAGAGAAATAGCAGCGGACTATATTGTAGC 4456
DB	3259 AAACATATTTTAAAGAGATATAGRAGAGAGAGAGAGCAGCGCTACAGATCTGTAGC 3318
QY	4457 CAGCTACAGCAGCGGCTCAAGACGTAATGTGTATATGACAGTAGGCACGATTA 4516
DB	3319 CAGCTGCAGCAGCGGCTCTAAGACGTAATGTGTG--TATGACAGTGAGGCCAAGTATTA 3376
QY	4517 AGAGTATAGTAAGCAACTATTTGATGAATTAGCTATTACATTAACTATGAT 4567
DB	3377 ATAGTATAGTAAGCAACTTTTGATGAATTGGCTATTGGCTTAGATGAT 3427
RESULT	6

PR	23-AUG-1999;	99US-0149902;
PR	23-AUG-1999;	99US-0149930;
PR	25-AUG-1999;	99US-0150566;
PR	26-AUG-1999;	99US-0150884;
PR	27-AUG-1999;	99US-0151065;
PR	27-AUG-1999;	99US-0151066;
PR	27-AUG-1999;	99US-0151080;
PR	30-AUG-1999;	99US-0151303;
PR	31-AUG-1999;	99US-0151438;
PR	01-SEP-1999;	99US-0151930;
PR	07-SEP-1999;	99US-0152363;
PR	10-SEP-1999;	99US-0153070;
PR	13-SEP-1999;	99US-0153758;
PR	15-SEP-1999;	99US-0154016;
PR	16-SEP-1999;	99US-0154039;
PR	20-SEP-1999;	99US-0154779;
PR	22-SEP-1999;	99US-0155139;
PR	23-SEP-1999;	99US-0155486;
PR	24-SEP-1999;	99US-0155659;
PR	28-SEP-1999;	99US-0156458;
PR	29-SEP-1999;	99US-0156596;
PR	04-OCT-1999;	99US-0157117;
PR	05-OCT-1999;	99US-0157753;
PR	06-OCT-1999;	99US-0157865;
PR	07-OCT-1999;	99US-0158029;
PR	08-OCT-1999;	99US-0158232;
PR	12-OCT-1999;	99US-0158369;
PR	13-OCT-1999;	99US-0159293;
PR	13-OCT-1999;	99US-0159294;
PR	13-OCT-1999;	99US-0159325;
PR	14-OCT-1999;	99US-0159329;
PR	14-OCT-1999;	99US-0159330;
PR	14-OCT-1999;	99US-0159331;
PR	14-OCT-1999;	99US-0159637;
PR	14-OCT-1999;	99US-0160814;
PR	18-OCT-1999;	99US-0159584;
PR	21-OCT-1999;	99US-0160741;
PR	21-OCT-1999;	99US-0160767;
PR	21-OCT-1999;	99US-0160768;
PR	21-OCT-1999;	99US-0160770;
PR	21-OCT-1999;	99US-0160817;
PR	21-OCT-1999;	99US-0160815;
PR	22-OCT-1999;	99US-0160980;
PR	22-OCT-1999;	99US-0160981;
PR	22-OCT-1999;	99US-0160989;
PR	25-OCT-1999;	99US-0161404;
PR	25-OCT-1999;	99US-0161405;
PR	25-OCT-1999;	99US-0161406;
PR	26-OCT-1999;	99US-0161359;
PR	26-OCT-1999;	99US-0161360;
PR	26-OCT-1999;	99US-0161361;
PR	28-OCT-1999;	99US-0161920;
PR	28-OCT-1999;	99US-0161992;
PR	28-OCT-1999;	99US-0161993;
PR	29-OCT-1999;	99US-0162142;

Query Match	2.4%;	score 113.6; DB
Best local similarity	55.8%;	pred. No 2.5e-15

Qy	1181	ATACCTGGATGAGGAGAAGGTGTGGCATGGCGAGAGATCGGCGCAGTGTGGCGGACGACGA	1240
Db	2920	AGACCTGAATAATGCGAAGAGAGAAGGTTTGATCATGACGGAGGTCTTTAATCAAC	2861
Qy	1241	TGGCCATCCCGGGGGGTACTTCGTTGGCGTGGCGCCAGCGCAOGGTGTCGCTCTCCCCCT	1300
Db	2860	GATCCATATTGTTCCCGTAGCTCTCGCGCTCATGATGTCCACACGTCACCTCTCTCCTT	2801
Qy	1301	GGTACCAAGAGACAGCCGATGGCGCCCGGTCTCGAAGGCGACCTCGAGCGCGCGGA	1360
Db	2800	GATACCAACACCGCGTTGATCTCTCCGCGCATTTCCCTATCTCTCCGTTCTCTTGA	2741
Qy	1361	CGAGATCGGAGTAAAGAGTTCGGCTTGGCCGACAGTCGGGCATCTGGTGGCGCGCAGG	1420

Db 2740 CCATCTCTCTGTACAGTGGCTTCCACGCTCCCACTCTTTATGCGCGTCCACGCGAAG 2681  
 QY 1421 CGCAGGGGACGAGCGCGATGACAGGAAGTGGCGGAGCGAAG-----GAGGGGGTTAG 1474  
 Db 2680 CGCAGCGGACCAACCGATCACAGCGAATCGTTTCCAGGCGATTCTTCCACGCGTTAG 2621  
 QY 1475 CGAAGGACATCCCGGGACCGACCGCGAGTGGCTGGAGTGGATGCGGTTGTGGACGG 1534  
 Db 2620 CGAACGCCATCTCCGGACCTACTCCACACACTTACTCTGTGTCAATGTCAAGCTGTAGTG 2561  
 QY 1535 GTCGTGGGCTCTCTCCATCGAGTGGCGGAGAGGAGGATGGAGGGGTGGGGG 1594  
 Db 2560 GTCGTGGTCTTCTCCACCGAGATCTGGGACAGGAGGATCGATGATGTTGGTG 2501  
 QY 1595 CGCACTCCGGGGACCAATCGCTCCCA 1622  
 Db 2500 CGCATTCGGTGAAGGATTTATCCCA 2473

## RESULT 7

ABN98574

ID ABN98574 standard; DNA; 949 BP.

XX AC ABN98574;

XX DT 01-AUG-2002 (first entry)

XX XX Arabidopsis thaliana expressed polynucleotide SEQ ID NO 342.

XX DE Arabidopsis thaliana; plant; insecticide; fungicide; transgenic; stress;

XX KW disease; crop; thale cress; tolerance factor; insect; pathogen;

XX KW nutrition; ds.

XX OS Arabidopsis thaliana.

XX XX US200203281-A1.

XX PN 21-FEB-2002.

XX PD 26-JAN-2001; 2001US-0770445.

XX PF 27-JAN-2000; 2000US-178472P.

XX PR (GOREL/) GORLACH J.

XX PA (ANYI/) AN Y.

XX PA (HAMI/) HAMILTON C M.

XX PA (PRIC/) PRICE J L.

XX PA (RAIN/) RAINES T M.

XX PA (YUYI/) YU Y.

XX PA (NAME/) NAMEAKA J G.

XX PA (PAGE/) PAGE A.

XX PA (MATH/) MATHAW A V.

XX PA (LEDE/) LEDFORD B L.

XX PA (WOBES/) WOESSNER J P.

XX PA (HAAS/) HAAS W D.

XX PA (GARC/) GARCIA C A.

XX PA (KRIC/) KRICKER M.

XX PA (SLAT/) SLATER T.

XX PA (DAVI/) DAVIS K R.

XX PA (ALLE/) ALLEN K.

XX PA (HOFF/) HOFFMAN N.

XX PA (HURB/) HURBAN P.

XX XX Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;

XX PI Rameaka JG, page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;

XX PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;

XX PI Hurban P;

XX DR WPI; 2002-400781/43.

XX XX New Arabidopsis thaliana nucleic acid for identifying homologous genes,

XX PT producing compositions that modulate the expression or function of its

PT encoded protein, and mapping functional regions of protein -  
 XX Claim 1; SEQ ID NO 342; 49pp + Sequence Listing; English.  
 XX The invention relates to an Arabidopsis thaliana nucleic acid (I)  
 CC comprising a sequence capable of hybridising under stringent conditions  
 CC to a sequence selected from any one of 999 sequences (ABN98233-ABN98233),  
 CC given in the specification or its fragment. A polypeptide (II) encoded by  
 CC (I), a transgenic plant (III) comprising an exogenous nucleic acid or a  
 CC genetically modified cell (IV) comprising an exogenous nucleic acid, is  
 CC useful for screening a candidate agent for its biological effect. (I) is  
 CC useful in identifying homologous or related genes, in producing  
 CC compositions that modulate the expression or function of its encoded  
 CC protein, mapping functional regions of the protein and in studying  
 CC associated physiological pathways. (I) is also useful for the genetic  
 CC manipulation of cells, particularly plant cells. (I) is also useful in  
 CC screening assays of various plant strains to determine the strains that  
 CC are best capable of withstanding a particular disease or environmental  
 CC stress. (II) and (III) are useful for screening of biologically active  
 CC agents, e.g. fungicides, insecticides, etc., for elucidating biochemical  
 CC pathways. The screened agents are useful in improved methods of treating  
 CC crops to prevent or treat disease. (II) are also useful in screening  
 CC programs to identify agents that mimic or enhance the action of tolerance  
 CC factors. Such agents are useful in improved methods of treating crops to  
 CC enhance their tolerance to environmental stress. (I) is also useful  
 CC for enhancing or inhibiting production of a biosynthetic product in a  
 CC plant. (III) is useful for identifying other mediators that may induce  
 CC expression of proteins of interest, for establishing the extent to which  
 CC a specific insect and/or pathogen is responsible for damage to a  
 CC particular plant, for identifying other mediators that enhance or induce  
 CC tolerance to environmental stress, for identifying factors involved in  
 CC biosynthetic pathways of nutritional, commercial, or medicinal value and  
 CC for identifying productions of nutritional, commercial or medicinal  
 CC value. (IV) is useful in the study of genetic function and regulation,  
 CC for alteration of the cellular metabolism and for screening compounds,  
 CC that may affect the biological function of the gene or gene products.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from USPTO  
 CC at seqdata.uspto.gov/sequence.html?docID=99990970445.

XX SQ Sequence 949 BP; 227 A; 248 C; 198 G; 276 T; 0 other;

Query Match 2.4%; Score 118.4; DB 24; Length 949;

Best Local Similarity 55.8%; Pred. No. 1.9e-15;

Matches 249; Conservative 0; Mismatches 191; Indels 6; Gaps 1;

QY 1183 ACCTGGATGAGGAGAGGTGTGGCATGCGGCGAGGTTCGGACGAGCATG 1242

Db 400 ACCTGANTATGGGAGAGAGAGAGGTGTGATGATGACGAGGTTCTTAATCAACGA 459

QY 1243 GCAATCCGGCGGGGTACTCTGTTGGCTCGGCCACGACGAGTGTGCTTCCCCCTGG 1302

Db 460 TCCATATTGTTCCGTTAGTCTCGGGTTCATGGATGTCCACACGTCACCTCTCTCTGA 519

QY 1303 TACCAGAGACAGCGCGATGCGGCGCGGTCTCGAGGGCGACCTTGGAGCGGGGACG 1362

Db 520 TACCACAAACACCGCTTGAATCTCTCGCGCGCAATTCCTACTCTCTCTCTTGTACC 579

QY 1363 AGATCGGAGTAGAGTCTGTCGCTTGGCCAGTGGCCATCTGTCGCGCGGCGGCG 1422

Db 580 ATCTCTCTGTACAGTGGCTTCACGCTCCCACTCTTTATGCGCGTTCACCGGAACGG 639

QY 1423 CAGGGGACGAGCGCGATGACAGGGAATCGCGGAGCGAAG-----GAGGGCGTTAGCG 1476

Db 640 CACGGCACCAACCGATCACAGCGGAATCTGTTCCAGGCGATTCTTACCGCGTTAGCG 699

QY 1477 AGGACATCCGCGGACCGACGCGAGTGTGAGTGTGATGCGCTTGTGGAGCGCG 1536

Db 700 AACGCCAATCTCTGGACCTACTCCACACACTTTTACCTGTGTCAATGTCAACGTGTAGTGGC 759

QY 1537 TCGTGGGCTCTCTCCATCGAGTCTCGGGGAGAGGCGGAGGATGGAGGGGTGGGGCG 1596

Db 760 TCGTGTGTTCTTCTCCACCGAGATCTCGGAGACAGCGGAGGATGATGAGTTTGTGCG 819

QY 1597 CACTCCGGGGGAACCATGCGGTCCCA 1622  
||| ||||| ||| ||| ||| |||  
Db 820 CACTCCGGTGGAGGATTTATCCCA 845

RESULT 8  
AAC45528/c  
ID AAC45528 standard; DNA; 1024 BP.  
XX AAC45528;  
XX AC AAC45528;  
XX DT 18-OCT-2000 (first entry)  
XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 46825.  
XX KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX OS Arabidopsis thaliana.  
XX PN EP1033405-k2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-0301439.  
XX PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127482.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0135124.  
PR 20-MAY-1999; 99US-0134941.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147152.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.

PR	13-AUG-1999;	99US-0148684.	
PR	16-AUG-1999;	99US-0149368.	
PR	17-AUG-1999;	99US-0149175.	
PR	18-AUG-1999;	99US-0149426.	
PR	20-AUG-1999;	99US-0149722.	
PR	20-AUG-1999;	99US-0149723.	
PR	20-AUG-1999;	99US-0149929.	
PR	23-AUG-1999;	99US-0149902.	
PR	23-AUG-1999;	99US-0149930.	
PR	25-AUG-1999;	99US-0150566.	
PR	26-AUG-1999;	99US-0150884.	
PR	27-AUG-1999;	99US-0151065.	
PR	27-AUG-1999;	99US-0151066.	
PR	27-AUG-1999;	99US-0151080.	
PR	30-AUG-1999;	99US-0151303.	
PR	31-AUG-1999;	99US-0151438.	
PR	01-SEP-1999;	99US-0151930.	
PR	07-SEP-1999;	99US-0152363.	
PR	10-SEP-1999;	99US-0153070.	
PR	13-SEP-1999;	99US-0153758.	
PR	15-SEP-1999;	99US-0154018.	
PR	16-SEP-1999;	99US-0154039.	
PR	20-SEP-1999;	99US-0154779.	
PR	22-SEP-1999;	99US-0155139.	
PR	23-SEP-1999;	99US-0155486.	
PR	24-SEP-1999;	99US-0155659.	
PR	28-SEP-1999;	99US-0156458.	
PR	29-SEP-1999;	99US-0156596.	
PR	04-OCT-1999;	99US-0157117.	
PR	05-OCT-1999;	99US-0157753.	
PR	06-OCT-1999;	99US-0157865.	
PR	07-OCT-1999;	99US-0158029.	
PR	08-OCT-1999;	99US-0158232.	
PR	12-OCT-1999;	99US-0158369.	
PR	13-OCT-1999;	99US-0159293.	
PR	13-OCT-1999;	99US-0159294.	
PR	13-OCT-1999;	99US-0159295.	
PR	14-OCT-1999;	99US-0159329.	
PR	14-OCT-1999;	99US-0159330.	
PR	14-OCT-1999;	99US-0159331.	
PR	14-OCT-1999;	99US-0159637.	
PR	14-OCT-1999;	99US-0159638.	
PR	18-OCT-1999;	99US-0159584.	
PR	21-OCT-1999;	99US-0160741.	
PR	21-OCT-1999;	99US-0160767.	
PR	21-OCT-1999;	99US-0160768.	
PR	21-OCT-1999;	99US-0160770.	
PR	21-OCT-1999;	99US-0160814.	
PR	21-OCT-1999;	99US-0160815.	
PR	22-OCT-1999;	99US-0160980.	
PR	22-OCT-1999;	99US-0160981.	
PR	22-OCT-1999;	99US-0160989.	
PR	25-OCT-1999;	99US-0161404.	
PR	25-OCT-1999;	99US-0161405.	
PR	25-OCT-1999;	99US-0161406.	
PR	26-OCT-1999;	99US-0161359.	
PR	26-OCT-1999;	99US-0161360.	
PR	26-OCT-1999;	99US-0161361.	
PR	28-OCT-1999;	99US-0161920.	
PR	28-OCT-1999;	99US-0161992.	
PR	28-OCT-1999;	99US-0161993.	
PR	29-OCT-1999;	99US-0162142.	
Query Match			2.4%; Score 118.4; DB 21; Length 1024;
Best Local Similarity			55.8%; Pred. No. 2e-15;
Matches 249; Conservative			0; Mismatches 191; Indels 6; Gaps 1;
Qy	1183	ACCTGATGAGGAGAAGTGTGGCATGGCCAGATCGCGCAGGTTCGGACAGCAATG	1242
Db	708	ACCTGATATATGGGAGAGAGAGGTTCAGATCATGACGGAGGTTCTTTAATCAACGA	649
Qy	1243	GCCATCGCGGGCGTACTCTGTGGCTGGCCCGCAGCGGTGTGCTCTCCCGCTGG	1302

Db	648	TCCATATTGTTCCCGTAGCTCTCGGGTCAATGGAATGCCAACAGTCACTCTCTCTGA	589
Qy	1303	TACCAAGAGACAGCCGATCGCGCGCGGTCCTCGAGGGCGACCTGGAGGGGGGACG	1362
Db	588	TACCAACACACCGCTTGATCTCTCGCGCGCATTTCTACTCTCTCTCTGACC	529
Qy	1363	AGATCGGATAGAGTGGTCCCTTTGGCCCACTGTCGGCCAGTCCCTGGTGGCGGACGGCG	1422
Db	528	ATCTCTCTGTAAGTGGCTTCCACGCTCCCACTCTTTTATCGCGTTCACCGGAAGCG	469
Qy	1423	CAGGGACAGCGCGATCAGACGAAGTGGCGGAGCGAAG-----GAGGGGTTAGCG	1476
Db	468	CACGGCACCAACCGGATCAGACCGCAATCTGTTCCAGGCGATCTTCACCGCGTATGCG	409
Qy	1477	AAGGACATCCCGGACCGCGCGCAGGTGCGGTTGGAGTCGATCGCGTTGTGGAGCGCG	1536
Db	408	AAGCCATCTCTGGACCTACTCCACACATTTACCTGTCAATGTCACTGTAGTGGC	349
Qy	1537	TGCTGGGCTCTCCCATCGAGCTCGCGGAGAGCGGAGGATGAGGGGTTGGGGCG	1596
Db	348	TCGTGTGCTTCTTCCACCGGAGATCTGCGGACAGCGGAGGATGATGATGTTGTGCG	289
Qy	1597	CATCGGGGGAACCATCGCGTCCCA	1622
Db	288	CATTCGCTGAGGATTTATCCCA	263
RESULT 9			
AAC46278/C			
ID	AAC46278 standard; DNA; 3108 BP.		
XX	AAC46278;		
DT	18-OCT-2000 (first entry)		
XX	Arabidopsis thaliana DNA fragment SEQ ID NO: 49559.		
DE	Arabidopsis thaliana.		
KW	Hybridisation assay; genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway;		
XX	metabolic pathway; promoter; termination sequence; ss.		
OS	Arabidopsis thaliana.		
XX	EP1033405-A2.		
PD	06-SEP-2000.		
XX	25-FEB-2000; 2000EP-0301439.		
PR	25-FEB-1999;	99US-0121825.	
PR	05-MAR-1999;	99US-0123180.	
PR	09-MAR-1999;	99US-0123548.	
PR	23-MAR-1999;	99US-0125788.	
PR	25-MAR-1999;	99US-0126264.	
PR	29-MAR-1999;	99US-0126785.	
PR	01-APR-1999;	99US-0127462.	
PR	06-APR-1999;	99US-0128234.	
PR	08-APR-1999;	99US-0128714.	
PR	16-APR-1999;	99US-0129845.	
PR	19-APR-1999;	99US-0130077.	
PR	21-APR-1999;	99US-0130449.	
PR	23-APR-1999;	99US-0130510.	
PR	28-APR-1999;	99US-0130891.	
PR	30-APR-1999;	99US-0131449.	
PR	30-APR-1999;	99US-0132048.	
PR	04-MAY-1999;	99US-0132484.	
PR	05-MAY-1999;	99US-0132485.	
PR	06-MAY-1999;	99US-0132486.	
PR	07-MAY-1999;	99US-0132487.	
PR	11-MAY-1999;	99US-0132863.	
PR	14-MAY-1999;	99US-0134256.	
PR	14-MAY-1999;	99US-0134218.	

PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134769.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144003.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144322.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 20-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155133.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.

[illegible]

PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145066.  
PR 21-JUL-1999; 99US-0145086.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145981.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.

PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160960.  
PR 22-OCT-1999; 99US-0160981.  
PR 25-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 26-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.  
PR 29-OCT-1999; 99US-0162142.

Query Match 2.4%; Score 118.4; DB 21; Length 3187;  
Best Local Similarity 55.8%; Pred. No. 3.le-15;  
Matches 249; Conservative 0; Mismatches 131; Indels 6; Gaps 1;  
QY 1183 ACCTGGATGAGGAGAGGTGTGCATGGCGGAGATCGCGCCAGGTGGCGGACGAGCATG 1242  
DB 2997 ACCTGAATATGGGAAGAGAGAGGTGTGAGATCATGACGAGGTTCTTAATCAACGA 2938  
QY 1243 GCCATCCGCGGGCGGTACTCGTTGGGTCGGCCCGACGCGACGGTGTCTCTCCCTGG 1302  
DB 2937 TCCATATTGTTCCCGTAGCTCTCGGCTCATGATGTCCACACGTCACCTCTCTCTGA 2878  
QY 1303 TACAGAGGACAGCCGATGCGGCGCGCGTCTCGAGGGGACCCCTGGAGCGGCGGAG 1362  
DB 2877 TACCACACACCGCCTTGATCTCTCGCGCGCATTTCTCTACTCTCTCTCTCTCTGACC 2818  
QY 1363 AGATCGGAGTAGAGTGTGGTGGCCCTTGGCCCATCTGTGGCCGATCTGTGGCCGCGG 1422  
DB 2817 ATCCTCTCGTACAGTGGCTTCCACGCTCCCACTCTTTATTCGCCGTTCACCGGAGCG 2758  
QY 1423 CAGGGACGAGCCGATGACAGGAACTGGCCGAGCGAAG-----GAGGGGTTAGCG 1476  
DB 2757 CAGGGACCAACCCGATCACAGCGAATCTGTTCCAGCGCATTTCTTCACCGGTTAGCG 2698  
QY 1477 AAGGACATCCCGGAGCAGCCGCGAGTGGGTTGGAGTGCATGCGCTGTGGACCGGC 1536  
DB 2697 AACGCCATTCCTGGACCTACTTCCACACACTTACCTGTGTCAATGTCAAGTGTAGTGGC 2638  
QY 1537 TCGTGGCCCTCTCCCATCGGAGCTGGGGGAGAGCGGAGGATGGAGGGTTGGGGCG 1596  
DB 2637 TCGTGTGCTTCTTCCACCGGAGATCTGGGACGAGCGGAGGATCGATGATGATGGTGG 2578  
QY 1597 CACTCCGGGGAAACCATGCGGTCCTCA 1522  
DB 2577 CATTCCGGTGAAGGATTTATCCCA 2552

RESULT 11  
AAC41895/c  
ID AAC41895 standard; DNA; 1024 BP.  
XX



AC AAC41895;  
XX 17-OCT-2000 (first entry)  
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 33532.  
XX  
XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX  
XX Arabidopsis thaliana.  
XX EP1033405-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126584.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 08-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145066.  
PR 21-JUL-1999; 99US-0145068.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 16-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149266.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 23-AUG-1999; 99US-0150566.

QY	1423	CAGGGGACGAGGCGCATGACAGGAA-----CTGGCCGGGAGGAGGCGCTTAGCG	1476
Db	468	CACGCGACCAACCGATGACAGCGCAATCTGTTTCCACGCAATCTTTCACGCGTTTAGCG	409
QY	1477	AAGGACATCCGGGACCGACGCGCAGGTGCGGTGGAGTCGATGCCGTTCTGGGAGCGGC	1536
Db	408	RACGCGATTCCTGGACCTACTCCACACATTTACTCTGTGCAATGTCAACGTGATGGC	349
QY	1537	TCGTGGGCGCTCCCTCCATCCGAGCTCGCGGGAGGCGGAGGATGAGGGGTTGGGGCGC	1596
Db	348	TCGTGTGCTTCTTCCACCGGAGATCTGGGACAGCGGAGGATGATGATGTTGTGGC	289
QY	1597	CACTCCGGGGACCATCGCTCCAG	1623
Db	288	CATTCCGCTGGGAGATTATCCAG	262
RESULT 12			
ID	AAAT9573/c		
XX	AAAT9573 standard; cDNA; 547 BP.		
AC	AAA79573;		
DT	27-NOV-2000 (first entry)		
XX	Pinus radiata cell signalling involved polynucleotide SEQ ID NO:374.		
XX	Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;		
KW	plant cell signalling; modulation; transgenic plant; pathogen; growth;		
KW	environmental change; development; cell proliferation; differentiation;		
KW	elongation; survival; disease resistance; nutrient metabolism; ss.		
XX	Pinus radiata.		
OS	Pinus radiata.		
XX	WO200042171-A1.		
PN	20-JUL-2000.		
PD	11-JAN-2000; 2000WO-US00724.		
XX	12-JAN-1999; 99US-0228986.		
XX	01-NOV-1999; 99US-0162866.		
XX	{GENE-} GENESIS RES & DEV CORP LTD.		
PA	Strabala IJ, Nieuwenhuizen NJ;		
PI	WPI; 2000-476052/41.		
DR	Isolated polynucleotide encoding a polypeptide involved in cell		
PT	signaling used for generating transgenic plants with modified responses		
PT	to external signals -		
XX	Claim 1; Page 176; 527pp; English.		
PS	AAAT9263 to AAAT9736 and AAAT95100 to AAAT9570 represent polynucleotide		
XX	and protein sequences isolated from eucalyptus (Eucalyptus grandis) or		
CC	pine (Pinus radiata also known as Monterey pine). The protein sequences		
CC	are involved in cell signalling. The polynucleotide and protein		
CC	sequences can be used to modify the response of plant cells to external		
CC	signals e.g. environmental changes or pathogens during the growth and		
CC	development of a plant. They can be used to modify cell proliferation,		
CC	differentiation, elongation and survival, resistance to disease and		
CC	nutrient metabolism. Examples of modifications which can be produced are		
CC	altered fruit ripening and senescence of leaves and flowers e.g. to		
CC	delay senescence and prolong the life of cut flowers or enhance		
CC	senescence of reproductive organs to engineer sterile plants. Other		
CC	modifications can be used to delay senescence in selected cell types or		
CC	organs providing fruit and vegetables which have a longer shelf life		
CC	between harvest and consumption, or to decrease branching frequency in		
CC	forest tree species giving long stretches of valuable knot-free clear		
CC	wood which can be used in solid timber furniture and veneers.		



The invention relates to a polynucleotide isolated from chromosome 11 of Indica rice cultivar CO39, flanked by marker R316 and RG1094 comprising one or more genes that confer resistance to strains of Magnaporthe grisea having avirulence gene AVR1-CO39. The Pi-CO39(t) polynucleotides are useful for conferring or improving resistance of plants to strains of the rice blast pathogen, Magnaporthe grisea and other pathogens. The Pi-CO39(t) nucleic acids may be used as probes to detect the presence of and/or expression of Pi-CO39(t) genes, and to produce large quantities of pure Pi-CO39(t) proteins. Purified gene products of Pi-CO39(t) are useful in producing polyclonal or monoclonal antibodies useful as sensitive detection reagents for the presence and accumulation of Pi-CO39(t) polypeptides. The transgenic plants are useful for plant breeding or directly in agricultural or horticultural applications. The present sequence is rice RGA38 contig.26Nippon DNA. This sequence comprises a portion of BAC clone 82N20 from Nipponbare, containing RGA38 sequence as well as ORFs NBR5 and NBR6.

```
Query Match          2.2%; Score 107.8; DB 20; Length 5526;
Best Local Similarity 78.1%; Pred. No. 7.3e-13;
Matches 143; Conservative 0; Mismatches 37; Indels 3; Gaps 1;
```

Seq	Sequence	15686 BP;	4645 A;	3184 C;	3204 G;	4653 T;	0 other;
Qy	Query Match	2.0%;	Score 99;	DB 24;	Length 15686;		
	Best Local Similarity	75.4%;	Pred. No. 8.6e-11;				
	Matches 150;	Conservative	0;	Mismatches 45;	Indels	4;	Gaps
Qy	4181	CAACGTTTGACCAITCGCTCTATTATTGAAAAAGATATATGATTTAGTATTTTATGTTATT	4240				
Db	9054	CCAACCTTTGATCGCCGCTTTATTTTGAGAAATATTTTATAATTAGTATTTTGTGTTATG	9113				
Qy	4241	AGATGNTAAACNTGAATAGTACTCTTTATGTCGAGCTAAATTTTTTTTATATATTTTTTATT	4300				
Db	9114	AGATGATTAATATATTAATAGTACTTTACTCAGACT--TATGTTTTTAATTTTTTTTCAAA	9171				
Qy	4301	AAATTTTCAAATAGACGGATGGTCNAAGCGCTAAACATGGATATCTATGGCTACACTTA	4360				
Db	9172	AAATTTTCAAATAGACAGACAGCTAATAA--GTTGGGCGCGGAAACAACTATGTTCACTTA	9229				

AAD38801  
ID AAD38801 standard; DNA: 15686 bp.

DT 23-SEP-2002 (first entry)

XX Rice RGA38 contig.26Nippon DNA.  
 DE  
 XX  
 KW Rice; chromosome 11; Indica rice cultivar; CO39; avirulence gene;  
 KW AVR1-CO39; rice blast pathogen; transgenic plant; plant breeding;  
 KW resistance: agricultural; horticultural; plant protectant; ds.

XX PN WO200234927-A2.

02-MAY-2002.

AA  
PR 20-OCT-2000; 2000US-242313P.

PA (WISC ) WISCONSIN ALUMNI RES FOUND.  
PA (USDA ) US DEPT OF AGRICULTURE.

DR WPI; 2002-471442/50.

PS Example 2; Page 52-59; 175pp; English.

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2003, 19:50:00 ; Search time 5095.83 Seconds  
(without alignments)  
15894.084 Million cell updates/sec

Title: US-09-702-134-7212\_COPY\_62300\_67300

Perfect score: 5001

Sequence: 1 aggggacactacattcca.....tgccgcggagtcgccccg 5001

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST: \*  
1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estm.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_htc.\*  
9: gb\_esti.\*  
10: gb\_est2.\*  
11: gb\_htc.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: gb\_gss.\*  
18: em\_gss\_hum.\*  
19: em\_gss\_inv.\*  
20: em\_gss\_pln.\*  
21: em\_gss\_vrt.\*  
22: em\_gss\_fun.\*  
23: em\_gss\_mam.\*  
24: em\_gss\_mus.\*  
25: em\_gss\_other.\*  
26: em\_gss\_pro.\*  
27: em\_gss\_tod.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
C 1	375.8	7.5	465	9	AU065751	AU065751	
C 2	334.4	6.7	498	9	AU030601	AU030601	
C 3	274.2	5.5	555	10	BE490225	BE490225 WHE0366_H	
C 4	254.6	5.1	309	14	D47552	D47552 RICS13122A	
C 5	230.4	4.6	381	14	BQ240809	BQ240809 TaE05012F	
C 6	210.6	4.2	486	17	AZ918792	AZ918792 1006008A0	

C 7	133	3.9	588	12	BF619879	BF619879 HVSMEC000
C 8	188.4	3.8	774	17	AQ576807	AQ576807 nbxb0089H
C 9	187.4	3.7	273	9	AU097543	AU097543 HVSME000
C 10	180	3.6	816	13	BI956651	BI956651 HVSME000
C 11	178.6	3.6	394	17	BH643146	BH643146 100805280
C 12	178.2	3.6	613	10	BE443619	BE443619 WHE1116_B
C 13	167	3.3	446	14	BQ237748	BQ237748 TaE05012F
C 14	164.8	3.3	535	17	AQ573895	AQ573895 nbxb0083A
C 15	163.6	3.3	842	17	AQ327645	AQ327645 nbxb0041B
C 16	162	3.2	589	17	AQ289925	AQ289925 nbxb0036E
C 17	161.6	3.2	689	17	AQ289077	AQ289077 nbxb0034K
C 18	161.4	3.2	502	12	BG560351	BG560351 RH12_73_
C 19	160.4	3.2	815	10	BE213308	BE213308 EST0065 T
C 20	160	3.2	579	17	AQ274081	AQ274081 nbxb0032M
C 21	157	3.1	757	17	AQ915616	AQ915616 nbxb0059C
C 22	153.2	3.1	600	17	AQ364616	AQ364616 nbxb0061C
C 23	153	3.1	423	12	BG102400	BG102400 RH12_23_
C 24	153	3.1	696	17	AQ689068	AQ689068 nbxb0078J
C 25	152.8	3.1	363	12	BF473797	BF473797 WHE0837_E
C 26	152.4	3.0	756	17	AQ689102	AQ689102 nbxb0033P
C 27	152.4	3.0	812	17	AZ135604	AZ135604 OSJNB0011
C 28	152.2	3.0	469	17	AQ446510	AQ446510 nbxb0070K
C 29	152.2	3.0	783	17	AQ863548	AQ863548 nbxb0021C
C 30	150.8	3.0	697	17	AQ579470	AQ579470 nbxb0084C
C 31	150.8	3.0	808	17	AQ794105	AQ794105 nbxb0052E
C 32	150.2	3.0	334	9	AL829579	AL829579 AL829579
C 33	146	2.9	610	17	AQ156838	AQ156838 nbxb0008D
C 34	145.8	2.9	665	17	AZ047724	AZ047724 nbxb0094A
C 35	144.4	2.9	720	17	AQ689393	AQ689393 nbxb0079S
C 36	144	2.9	583	14	BQ253395	BQ253395 san69f02_
C 37	141.8	2.8	667	17	AQ864314	AQ864314 nbxb0022N
C 38	141.8	2.8	763	17	AQ914146	AQ914146 nbxb0047M
C 39	140.8	2.8	735	17	AQ864484	AQ864484 nbxb0023I
C 40	140.8	2.8	847	17	AQ868460	AQ868460 nbxb0027J
C 41	139.8	2.8	657	17	AZ131579	AZ131579 OSJNB0011
C 42	139.6	2.8	554	13	B1128518	B1128518 G091P77Y
C 43	138.8	2.8	764	17	AQ914191	AQ914191 nbxb0047H
C 44	138.2	2.8	456	13	BF643717	BF643717 IPL 56-80
C 45	136	2.8	578	17	AQ509468	AQ509468 nbxb0096I

#### ALIGNMENTS

RESULT 1	AU065751/c	465 bp	mRNA	linear	EST 02-APR-2002
LOCUS	AU065751	Rice cDNA from immature leaf including apical meristem			
DEFINITION	Oryza sativa (japonica cultivar-group) cDNA clone B51292_1A, mRNA sequence.				
ACCESSION	AU065751				
VERSION	AU065751.1	GI:4968847			
KEYWORDS	EST.				
SOURCE	Oryza sativa (japonica cultivar-group).				
ORGANISM	Oryza sativa (japonica cultivar-group).				
REFERENCE	1 (bases 1 to 465)				
AUTHORS	Sasaki,T. and Yamamoto,K.				
TITLE	Rice cDNA from immature leaf including apical meristem				
JOURNAL	Unpublished (1997)				
COMMENT	Contact: Takuji Sasaki National Institute of Agrobiological Resources Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan Tel: 81-298-38-7441 Fax: 81-298-38-7468 Email: tsasaki@abrr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/ PROJECT = "RGP" Location/Qualifiers 1. .465 /organism="Oryza sativa (japonica cultivar-group)"				

```

/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="E51292_1A"
/clone_lib="Rice cDNA from immature leaf including apical meristem"
/dev_stage="immature"
/notes="Organ: leaf; immature leaf including apical meristem (under long day condition)"
BASE COUNT 72 a 160 c 138 g 90 t 5 others
ORIGIN

```

```

Query Match 7.5%; Score 375.8; DB 9; Length 465;
Best Local Similarity 98.5%; Pred. No. 1e-57;
Matches 388; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1426 GGCACGAGCCCATCAGAGGAAGTGGCCGAGCGAGCGAGGAGGCGCTTAGCGAGGACATC 1485
|||||
DB 393 GGGACCAAGCCGATWACAGGNACTGGCCGGAGCGAAGGA-GGGCTTAGCGAGGACATC 335
|||||
QY 1486 CCGGGACCGACCGCCAGGTGGGTGGAGTCGATGCCGTTGTGGAGCGCTCGTGGGCC 1545
|||||
DB 334 CCGGGACCGACCGCGAGTGGGTGGAGTCGATGCCGTTGTGGAGCGCTCGTGGGCC 275
|||||
QY 1546 TCCTCCATCGAGCTCGGGGAGAGCGGAGGATGGAGGGTTGGGGGCGCACTCCGGG 1605
|||||
DB 274 TCCTCCATCGAGCTCGGGGAGAGCGGAGGATGGAGGGTTGGGGGCGCACTCCGGG 215
|||||
QY 1506 GGAACCATCGCTCCAGTGGACCCACACCGCCCGCCGCGCCGCGCATGTTCCGACTGC 1665
|||||
DB 214 GGAACCATCGCTCCAGTGGACCCACACCGCCCGCCGCGCCGCGCATGTTCCGACTGC 155
|||||
QY 1666 CTCCGAGGATGAACACCACTGTTTCGAGGAGGAGGAGACAGCGCCGCGCGCGCCG 1725
|||||
DB 154 CTCCGAGGATGAACACCACTGTTTCGAGGAGGAGGAGACAGCGCGCGCGCGCGCCG 95
|||||
QY 1726 GGAAGCAGCAGCAAGCAGCAGCATCCGCGCATGCTGCTTCTACGCTTGGGCTGG 1785
|||||
DB 94 GGAAGCAGCAGCAAGCAGCAGCATCCGCGCATGCTGCTTCTACGCTTGGGCTGG 35
|||||
QY 1786 GGCTCGGGGGTGGGTTTGGGCAATAATTG 1819
|||||
DB 34 GGCTCGGGGGTGGGTTTGGGCAATAATTG 1
|||||

```

```

RESULT 2
AU030601/c 498 bp mRNA linear EST 01-APR-2002
LOCUS AU030601.1
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone E51292_2L, mRNA sequence.
ACCESSION AU030601
VERSION AU030601.1 GI:3763862
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group).
ORGANISM Oryza sativa (japonica cultivar-group).
REFERENCE 1
AUTHORS Sasaki,T. and Yamamoto,K.
TITLE Rice cDNA from immature leaf including apical meristem
JOURNAL Unpublished (1997)
COMMENT Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@nri.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
PROJECT = "RGP".
POLYA=No.
FEATURES
source Location/Qualifiers
1..498

```

```

/organism="Oryza sativa (japonica cultivar-group)"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="E51292_2Z"
/clone_lib="Rice cDNA from immature leaf including apical meristem"
/dev_stage="immature"
/notes="Organ: leaf; immature leaf including apical meristem (under long day condition)"
BASE COUNT 128 a 125 c 118 g 127 t
ORIGIN

```

```

Query Match 6.7%; Score 334.4; DB 9; Length 498;
Best Local Similarity 96.9%; Pred. No. 2.9e-50;
Matches 341; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 98 TTTTAATTATACCAAGCCCAATTCATATATCAAGAAATGAGTGGTCCATGAACTAAATA 157
|||||
DB 487 TTTTATATACCAAGCCCAATTCATATATCAAGAAATGAGTGGTCCATGAACTAAATA 428
|||||
QY 158 GAAAGTAGACATAGAGTGGTGGTCTTCTAGCTATTATCTGATATAAAGAAACACAC 217
|||||
DB 427 GAAAGTAGACATAGAGTGGTGGTCTTCTAGCTATTATCTGATATAAAGAAACACAC 368
|||||
QY 218 ACAGAGAGACTAGAGTGGTGGTGGTCTTCAATATAGCTTGGAGCCATGTCATATGTC 277
|||||
DB 367 ACAGAGAGACTAGAGTGGTGGTGGTCTTCAATATAGCTTGGAGCCATGTCATATGTC 308
|||||
QY 278 CAAGCTGAGCTGGTGGTGGTGGTCTTCAATATAGCTTGGAGCCATGTCATATGTC 337
|||||
DB 307 CAAGCTGAGCTGGTGGTGGTGGTCTTCAATATAGCTTGGAGCCATGTCATATGTC 248
|||||
QY 338 TTGCATCAACAAATCTCACGTTTCGAGTTTGTATCCCTTCTGAGCTTCCCTTACTACTT 397
|||||
DB 247 TTGCATCAACAAATCTCACGTTTCGAGTTTGTATCCCTTCTGAGCTTCCCTTACTACTT 188
|||||
QY 398 CAGTATATGTCAGAGCCCTGATGCTAGTCCCTGATGAAACCAACACAG 449
|||||
DB 187 CAGTATATGTCAGAGCCCTGATGCTAGTCCCTGATGAAACCAACAG 136
|||||

```

```

RESULT 3
BE490225/c 555 bp mRNA linear EST 31-JUL-2000
LOCUS BE490225.1
DEFINITION Wheat cold-stressed seedling cDNA library
Triticum aestivum cDNA clone WHE0366_H01_P02, mRNA sequence.
ACCESSION BE490225
VERSION BE490225.1 GI:9609758
KEYWORDS EST.
SOURCE Bread wheat.
ORGANISM Triticum aestivum
REFERENCE 1
AUTHORS Anderson,O.D., Choi,S., Choi,D.W., Close,T.J., Penton,R.D., Han,P.S., Hsiao,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J., Seaton,C.L. and Tong,J.C.
TITLE The structure and function of the expressed portion of the wheat genomes - Cold-stressed seedling cDNA library
JOURNAL Unpublished (2000)
COMMENT Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@w.usda.gov
Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20
Seq primer: Stratagene SK primer.
FEATURES
source Location/Qualifiers
1..555

```

```

/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/gclone="WHE0366_H01_P02"
/gclone_lib="Wheat cold-stressed seedling cDNA library"
/tissue_type="Seedling"
/dev_stage="Five-day old seedling"
/lab_host="E. coli 501R"
/note="Vector: lambda Uni-ZAP XR, excised phagemid;
Site_1: EcoRI; Site_2: XhoI; Seeds were surface-sterilized
, germinated and grown aseptically in the dark at room
temperature on filter paper with water, nystatin and
cefotaxime in covered crystallization dishes. Five-day
old seedlings were transferred to 5 C cold room and kept
for 48 hr. The tissue, total RNA, and poly(A) RNA were
prepared, a cDNA library was made, and the cDNA clones
were in vivo excised to give phagescript phagemids in the
TJ Clonase lab (Choi, Clonase, Fenton) at the University of
California, Riverside. Plasmid DNA preparations and DNA
sequencing were performed in the OD Anderson lab (all
other authors)."
96 a 193 c 203 g 83 t

```

```

Query Match          5.5%; Score 274.2; DB 10; Length 555;
Best local Similarity 80.0%; Pred. No. 2.1e-39;
Matches 337; Conservative 0; Mismatches 78; Indels 6; Gaps 1;

QY 1183 ACCTGGATGAGGAGAGGTGTGGCATGCGAGATCGCGCGCAGGTGCGGACGACGATG 1242
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 421 ACCTGGATGAGGAGAGGTGTGGCATGCGGAGTCTCTCCGAGGTCCGAGACGAGGGCG 362

QY 1243 GCCATFCGGGCGCGCTACTCTGTGTGGCGTTCGCGCCACGCAAGGTGTGCTCTCCCGCTGG 1302
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 361 CCCATGTGGGCGCGCTACTCTGTGTGGCGTTCGCGCCACGCAAGGTGTGCTCTCCCGCTGG 302

QY 1303 TACCAGAGGACAGCGCCGATGCGCGCGGTCTCCAGGCGCACCTCGAGCGCGGACG 1362
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 301 TACCACAGACAGCGCCCGGATCGCGCGCGGTCTCCAGGCGCACCTCGAGCGCGGACG 242

QY 1363 AGATCGGAGTAGAGGTCTGTCCTCTTGGGCCACTCTGTGTCCGCGGACGGCG 1422
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 241 ATGTCTCGCTACAGTCTCGACACCTTTCGCCACTTCGCGCCATCCGCTGTCGCCGAGGGCG 182

QY 1423 CAGGGGACGAGGCCGATGACAG- ----GGAACTGCCGCGAGCGACGAGGGGGTTAGCG 1476
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 181 CAGGGCACAGGCGCGACGCGCGCGCGCGCGCGCGCGCGCGCGAGCGAGCGGTGTGGCG 122

QY 1477 AAGACATCTCCGGGACCGACGCGCGAGTGCGGTTGTGAGTCTGATCGCGTTGTGGAGCGCG 1536
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 121 AAGGACATCTCCGGGCGCGACCGCGAGTGCGGTTGTGCGCTGATGCCCTGTGCACGCGC 62

QY 1537 TCGTGGGCGCTCTCCCATCGAGTCTGGGGGACGAGCGGAGAGTGAAGGGTTGGGGCG 1596
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 61 TCAATGCGCTGCTCTCCACCGCAACAGGGGGAGAGGCGGAGCGCGCGCGCGCGCGCG 2

QY 1597 C 1597
|
Db 1 C 1

```

RESULT 4

LOCUS	D47552				
DEFINITION	D47552	309 bp	mRNA	linear	EST 03-APR-2002
	RIC53122A	Rice green shoot	Oryza sativa	(japonica cultivar-group)	
		CDNA, mRNA sequence.			
ACCESSION	D47552				
VERSION	D47552.1	GI:701261			
KEYWORDS	EST.				
SOURCE	Oryza sativa	(japonica cultivar-group).			
ORGANISM	Oryza sativa	(japonica cultivar-group)			
	Eukaryota;	Viridiplantae;	Streptophyta;	Embryophyta;	Tracheophyta;
	Spermatophyta;	Magnoliophyta;	Liliopsida;	Poales;	Poaceae;

Birhartoideae; Oryzeae; Oryza.  
1 (bases 1 to 309)  
AUTHORS Sasaki, T., Miyao, A. and Yamamoto, K.  
TITLE Rice cDNA from callus 1995  
JOURNAL Unpublished (1995)  
COMMENT Contact: Takuji Sasaki  
National Institute of Agrobiological Resources  
Rice Genome Research Program, Ramondal 2-1-2, Tsukuba, Ibaraki  
305-8602, Japan  
Tel: 81-298-38-7441  
Fax: 81-298-38-7468  
Email: tsasaki@agr.affrc.go.jp, URL: http://irp.dna.affrc.go.jp/  
PROJECT = 'RGP',  
Location/Qualifiers  
1. .309  
/organism="Oryza sativa (japonica cultivar-group)"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone\_lib="Rice green shoot"  
/note="green shoot (8 days old)"  
57 a 94 c 91 g 60 t 7 others  
BASE COUNT  
ORIGIN

Query Match 5.1%; Score 254.6; DB 14; Length 309;  
Best Local Similarity 93.9%; Pred. No. 8e-36;  
Matches 293; Conservative 0; Mismatches 16; Indels 3; Gaps 3;

QY 4682 CCTGTCCGTCACCTTCTGCTGTTAAATGCGTTCTAGGTAGCGTAGAGATCCAACAA 4741  
|||||  
Db 1 CCGTCCGTCACCTTCTGCTGTTAAATGCGTTCTAGGTAGCGTAGAGATCCAACAA 59  
|||||

QY 4742 ACACCTAGCAGCAGCAGCATTGGTACGTACTCTCTAGTCTCTAGTGTAGTGAGGTGAGCGT 4801  
|||||  
Db 60 ACACCTAGCAGCAGCAGCATTGGTACGTACTCTCTAGTGTAGTGAGGTGAGCGT 119  
|||||

QY 4802 GAGGAGCAGTGCAGTATGCGGTGGCGACAGTACAGTGCAGGTACTAGCGCTCTCTCTGC 4861  
|||||  
Db 120 GAGGAGCAGTGCAGTATGCGGTGGCGACAGTACAGTGCAGGTACTAGCGCTCTCTCTGC 178  
|||||

QY 4862 TGCTGGCCACACGCGCCAGCGGCGAGCGCGACGCGTGTATATCTCTTTTGG 4921  
|||||  
Db 179 TGCTGGTACACGCGCCAGCGGCGAGCGCGACGCGTGTATATCTCTTTTGG 238  
|||||

QY 4922 GCGGCGAGTCCACATGCGGTGTGCGGCGCGCCGCCACCAACGCGCGTGGAGCGGCGTGG 4981  
|||||  
Db 239 GCGGCGAGTCCACATGCGGTGTGCGGCGCGCCGCCACCAACGCGCGTGGAGCGGCGTGG 297  
|||||

QY 4982 TCGCGCGCGAGT 4993  
|||||  
Db 298 TCGCGCGCGAGT 309  
|||||

RESULT 5  
B0240809/c  
LOCUS B0240809 381 bp mRNA linear EST 03-MAY-2002  
DEFINITION TaB05012F06R TaE05 Triticum aestivum cDNA clone TaE05012F06R, mRNA  
sequence.  
ACCESSION B0240809  
VERSION B0240809.1 GI:20436685  
KEYWORDS EST.  
SOURCE bread wheat.  
ORGANISM Triticum aestivum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
; Triticeae; Triticum.  
1 (bases 1 to 381)  
Cloutier, S.  
Wheat functional genomics - Glenlea developing seeds cDNA libraries  
Unpublished (2002)  
Contact: Dr. Sylvie Cloutier  
Cereal Research Centre, Agriculture and Agri-food Canada  
195 Dafoe Rd, Winnipeg, MB, Canada R3T 2M9  
Tel: (204) 983-2340

Fax: (204) 983-4604  
 Email: scloutier@agr.ca  
 was cloned directionally, not all sequences generated with reverse  
 primer were from the 5' end (same with forward primer and 3' end).  
 Average insert size is >2.0 kb  
 Plate: 012 row: F column: 06  
 Seq primer: M13 Reverse.

## FEATURES

Location/Qualifiers

source

1. .381  
 /organism="Triticum aestivum"  
 /cultivar="Glenlea"  
 /db\_xref="taxon:4565"  
 /clone="TaE05012F06R"  
 /tissue\_type="developing seeds"  
 /dev\_stage="5 days after anthesis"  
 /lab\_host="E. coli DH10B"  
 /notes="Vector: pSPORT-P (Invitrogen Technologies); Site\_1:  
 NOL1; Site\_2: MluI; mRNA obtained from wheat seeds of  
 cultivar Glenlea 5 days post-anthesis"  
 39 a 140 c 156 g 46 t

BASE COUNT

ORIGIN

Query Match 4.6%; Score 230.4; DB 14; Length 381;  
 Best Local Similarity 78.1%; Pred. No. 1.8e-31;  
 Matches 292; Conservative 0; Mismatches 76; Indels 6; Gaps 1;  
 QY 1272 GGCACAGCGCAGGTGTCTCCCTCCCTGTACAGAGGACAGCGCGATCGCGCCGCC 1331  
 Db 381 GGCACAGCGCAGGTGTCTCCCTCCCTGTACAGAGGACAGCGCGATCGCGCCGCC 322  
 QY 1332 GGTCTGAGGGGACCTTGAGCGCGGACGAGATCGAGTAGAGTCGCTGCCCTTGGC 1391  
 Db 321 GGTCTCCAGCGACCGCGCGCGCGCCGACCATGTCGTCGACAGCTCCGACCTTGGC 262  
 QY 1392 CCAGTGGCCATCTCTCCGCGCGCGCGAGGCGCGATGACAGGGAATG 1451  
 Db 261 CCAGTGGCCATCTCTCCGCGCGCGCGAGGCGCGATGACAGGGAATG 202  
 QY 1452 -----GCGGAGGAGGAGGCGCTTACGAGGACATCCCGGGACCGCGCGAGGT 1505  
 Db 201 CGCGCGCGCGGAGGAGGAGGCGCTTGGCGAGGACATCCCGGGCGCGCGAGGT 142  
 QY 1506 GCGGTGTGAGTGCATCGCTGTGGAGCGGCTCGTGGGCTCCCTCCATCGGAGCTGGG 1565  
 Db 141 GCGGTGTGAGTGCATCGCTGTGGAGCGGCTCGTGGGCTCCCTCCATCGGAGCTGGG 82  
 QY 1566 GGAGGCGGAGGATGGAGGGGTTGGGGGCGACATCCCGGGGGAACCATGCTCCCATGTG 1625  
 Db 81 GGAGGCGGAGGATGGGCGCGCGAGGCGCGAGTCCCGGGCGCACCATCCCGCTCCCATGTG 22  
 QY 1626 GGACCCACACGC 1639  
 Db 21 CGTGGCGCGGAGCG 8

RESULT 6  
 AZ918792/c 486 bp DNA linear GSS 17-DEC-2001  
 LOCUS 1005008A02.y1 1006 - RescueMu Grid G Zea mays genomic, DNA  
 DEFINITION sequence.  
 ACCESSION AZ918792  
 VERSION AZ918792.1 GI:13388076  
 KEYWORDS GSS.  
 SOURCE Zea mays.

## ORGANISM

Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoideae; Andropogoneae; Zea.  
 1 (bases 1 to 486)  
 Walbot.V.  
 TITLE Maize genomic sequences found using engineered RescueMu transposon  
 JOURNAL Unpublished (2001)

## COMMENT

Contact: Walbot V  
 Department of Biological Sciences  
 Stanford University  
 855 California Ave, Palo Alto, CA 94304, USA  
 Tel: 650 723 2227  
 Fax: 650 723 8221  
 Email: walbot@stanford.edu

Sequence was trimmed at very probable ligation site. Post-ligation  
 sequence submitted separately.  
 Plate: 1006008 row: 3  
 Class: transposon-tagged

## FEATURES

Location/Qualifiers

Source

1. .486  
 /organism="Zea mays"  
 /cultivar="mixed background W23/A188/B73"  
 /db\_xref="taxon:4577"  
 /clone\_lib="1006 - RescueMu Grid G"  
 /tissue\_type="leaf"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /notes="Organ: leaf; Vector: RescueMu (engineered from  
 pBlueScript backbone); Site\_1: BamHI; Site\_2: BglII;  
 RescueMu is a 4.9 kb, modified maize Mu transposon  
 designed to allow plasmid rescue from total genomic DNA.  
 Mu elements insert preferentially into transcription  
 units. For more information on RescueMu, go to the web  
 site 'www.zmndb.iastate.edu' and follow the links for  
 'RescueMu.' Grid G was grown at Stanford in 2000. DNA was  
 extracted from leaf punches, double digested using BamHI  
 and BglII, and ligated to form circular plasmids. DH10B  
 cells were transformed and then screened on LB plates with  
 ampicillin."

BASE COUNT 57 a 176 c 181 g 72 t

ORIGIN

Query Match 4.2%; Score 210.6; DB 17; Length 486;  
 Best Local Similarity 74.2%; Pred. No. 6.3e-28;  
 Matches 296; Conservative 0; Mismatches 94; Indels 9; Gaps 2;  
 QY 1315 GCGCGATGCGCGCGCGCTCTCGAGGCGGACCTTGGAGCGCGGCGAGCATCGAGTAG 1374  
 Db 483 GCACCGATGCGCGCGCGCGCTCTCGAGGCGGACCTTGGAGCGCGGCGAGCATCGAGTAG 424  
 QY 1375 AGTGTGTCCTTGGCCCTGCGCCAGTCGCCATCTGTCGCGCGCGCGCGAGGAGG 1434  
 Db 423 AGTGTGTCCTTGGCCCTGCGCCAGTCGCCATCTGTCGCGCGCGCGAGGAGGAGG 364  
 QY 1435 CCGATGACAGGGAAC-----TGGCGCGAGCGAGGAGGCGGTAGCGAAGGACATCCG 1488  
 Db 363 CCGATGACAGGGAAC-----TGGCGCGAGCGAGGAGGCGGTAGCGAAGGACATCCG 304  
 QY 1489 GGACCGACCGCGCAGGTGCGGT---GGAGTCGATGCGGTGTGGAGCGGCTCGTGGGCC 1545  
 Db 303 GGCGCCACCGCGCAGGTGCGGTGTGGAGTCGATGCGCGGCTGCGAGCGGCTCGCGGCC 244  
 QY 1546 TCCTCCATCGAGTCGCGGGGAGGCGGAGGATGGAGGGTTGGGGCGGCACTCCGG 1605  
 Db 243 TCCTCCATCGAGTCGCGGGGAGGCGGAGGATGGAGGGTTGGGGCGGCACTCCGG 184  
 QY 1606 GGAACCATGCGCTCCAGTGGGACCCACACGCGCGCGCGCGCGCGCGCGCGCGCGCG 1665  
 Db 183 GGCACAAAAGTCCAGCGGTTCCGCCACCCCGCGCGCGCGCGCGCGCGCGCGCGCG 124  
 QY 1666 CTTCCGAGGTGACACACCATCTGTTTCGAGCGAGGAGG 1704  
 Db 123 CCGCGCAGGATGAACACCATCTGTTGGACGCGGCGCAG 85

## RESULT 7

BF619879/c

LOCUS

DEFINITION

BF619879 588 bp mRNA linear EST 22-OCT-2001  
 HVSMEC009122f Hordeum vulgare seedling shoot EST library  
 HVCNDR0003 (Etiolated and unstressed) Hordeum vulgare cDNA clone



HVSMEC0009122f, mRNA sequence.  
 BF619879  
 VERSION BF619879.2 GI:13107874  
 EST.  
 SOURCE Hordeum vulgare.  
 ORGANISM Hordeum vulgare.  
 REFERENCE 1 (bases 1 to 588)  
 AUTHORS Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D., Yu  
 ,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Choi,D.W., Fenton  
 ,R.D., Oates,R. and Main,D.  
 TITLE Development of a genetically and physically anchored EST resource  
 for barley genomics: Morex unstressed seedling shoot cDNA library  
 JOURNAL Unpublished (2001)  
 COMMENT On Dec 18, 2000 this sequence version replaced gi:11883613.  
 Contact: Wing RA  
 Clemson University Genomics Institute  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Tel: 864 656 7288  
 Fax: 864 656 4293  
 Email: rwing@clemson.edu  
 Total hg bases = 342  
 Seq primer: AATTAACTCCCTACTAAAGG  
 High quality sequence stop: 583.  
 Location/Qualifiers  
 1..588  
 /organism="Hordeum vulgare"  
 /cultivar="Morex"  
 /db\_xref="taxon:4513"  
 /clone="HVSMEC0009122f"  
 /glone\_lib="Hordeum vulgare seedling shoot EST library  
 HVCDNA0003 (Etiolated and unstressed)"  
 /issue\_type="Seedling shoot"  
 /lab\_host="TJ121"  
 /note="Vector: lambdaZAP; Site\_1: EcoRI; Site\_2: XhoI;  
 Seeds were surface sterilized then germinated under axenic  
 conditions in the dark at room temperature on filter paper  
 with water, nystatin and cefotaxime in covered  
 crystallization dishes. Five-day old seedling shoots were  
 then harvested, total RNA was prepared, poly(A) RNA was  
 purified, one primary unamplified cDNA library was made,  
 and 1 million pfu were in vivo excised to give pluescript  
 SK(-) cDNA phagemids. These steps were performed in the Yu  
 Close laboratory at the University of California,  
 Riverside (Choi, Close, Fenton). Phagemids were plated and  
 picked at the Clemson University Genomics Institute (CUGI)  
 (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA  
 preparations, DNA sequencing and sequence analysis were  
 performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates  
 , Rambo, Main). The sequence has been trimmed to remove  
 vector sequence and contains a minimum of 100 bases of  
 phred value 20 or above. For more details on library  
 preparation and sequence analysis see  
 http://www.genome.clemson.edu/projects/barley. To order  
 this clone see http://www.genome.clemson.edu/orders Also  
 see Close T.J., Wing R., Kleinhofs A., Wise R (2001)  
 Genetically and physically anchored EST resources for  
 barley genomics. Barley Genetics Newsletter 31:29-30.  
 (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"  
 91 a 198 c 203 g 96 t  
 BASE COUNT  
 ORIGIN  
 Query Match 3.9%; Score 193; DB 12; Length 588;  
 Best Local Similarity 62.7%; Pred. No. 9e-25;  
 Matches 339; Conservative 0; Mismatches 190; Indels 12; Gaps 2;  
 QY 1183 ACCTGATGAGGAGGTGTGCGAGATCGCGCGCAGGTGCGGACGAGCATG 1242  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 575 ACCTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 515

QY 1243 GCATCCGCGCGGCGTACTGTTGGGTGCGGCCAGCCAGCGATGCTCTCTCCCTGG 1302  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 515 TCCACGTTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 456  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 1303 TACACAGACAGCCCGATGCGCGCGCGCGCTCGAGGGGCGCCCTGGAGCGGCGAG 1362  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 455 TACACAGACAGCCCGATGCGCGCGCGCGCTCGAGGGGCGCGCGCGCGCGCGAG 399  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 1363 AGATCGAGTACAGATGCGGTGCGCTGGCGCCAGTGGCCATCTCTGCTGCGCGAGCGG 1422  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 398 ATCTGCTCTACAGTGTCTCCCGCGCGCGCCACTCCCGATGCGCGCTGCGCGCAAGCGG 339  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 1423 CAGGGAGCAG-----GCCGATGACAGAGAACTGCGCGGAGCGAGAGAGCGGTTA 1473  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 338 CACGCTACAGAGCCCTACCGCGCGGTGCTGCGCGCGGTGCTGCGCGGAGAAATGCGCGG 279  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 1474 GCGAAGGACATCCCGGAGCGAGCGCGAGTGGCTGTTGGAGTGCATGCGCTGTGTGAGC 1533  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 278 GCGAAGGACATCCCGGAGCGAGCGCGAGTGGCTGTTGGAGTGCATGCGCTGTGTGAGC 219  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 1534 GGCTGCTGGGCGCTCTCCATCGAGTGGAGTGGCGGAGAGCGCGAGAGTGGAGGTTGGGG 1593  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 218 GGCTGCTGGGCGCTCTCCATCGAGTGGAGTGGCGGAGAGCGAGAGTGGAGGTTGGGG 159  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 1594 GCGCACTCCGGGGAGCACTGCTCCAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 1653  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 158 GCGCACTCCGGGGAGCACTGCTCCAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 99  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 1654 ATGTTGACTGCTCCCGAGGATGACACACACCTGTTTCGACGAGGAGGAGGAGACAGCC 1713  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 98 ATGTTGCTGCTGCGCTGACAGAGAGAGATGCGCATGACTGCGGCGCTGCGGAGACAGCGC 39  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 1714 G 1714  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 38 G 38  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 RESULT 8  
 AQ576807/c 774 bp DNA linear GSS 02-JUN-1999  
 LOCUS nbxb0089H04r CUGI Rice BAC Library Oryza sativa genomic clone  
 DEFINITION nbxb0089H04r, DNA sequence.  
 ACCESSION AQ576807  
 VERSION AQ576807.1 GI:4977292  
 KEYWORDS GSS.  
 SOURCE Oryza sativa.  
 ORGANISM Oryza sativa.  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 REFERENCE 1 (bases 1 to 774)  
 AUTHORS Wing,R.A. and Dean,R.A.  
 TITLE A BAC End Sequencing Framework to Sequence the Rice Genome  
 JOURNAL Unpublished (1998)  
 COMMENT Contact: Wing RA  
 Clemson University Genomics Institute  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Tel: 864 656 7288  
 Fax: 864 656 4293  
 Email: rwing@clemson.edu  
 Seq primer: GGAACAGCTATGACCATG  
 Class: BAC ends  
 High quality sequence stop: 465.  
 Location/Qualifiers  
 1..774  
 /organism="Oryza sativa"  
 /strain="Japonica"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:4530"  
 /clone="nbxb0089H04r"  
 /clone\_lib="CUGI Rice BAC Library"  
 /tissue\_type="Leaf"

genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from *Oryza sativa*, Nipponbare variety. The library contains 36,864 clones with an average insert size of 138.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."

Query Match 3.7%; Score 187.4; DB 9; Length 273;  
Best Local Similarity 85.3%; Pred. No. 1.1e-23;  
Matches 191; Conservative 0; Mismatches 33; Indels 0; Gaps

QY 98 TTTAATTATACCAAGCCAAATTCAATAATCATAGAATGAGTGGTCCATGAAACTAA 157

224  
pH

[illegible]

158 GAAAGTAGAACATAAGGTGCGTAGTCTCTTAGCTATTATCTGATAAAAGAAAACAAC 217

Db 164 GAAAGNNGACAAAGGGCTAGNCNCTTAGNTATTATTCANGATAAAAGNPAACAAC 105

[illegible][illegible]

104 ACAGNGGNGCTAGAGCCNGGAAAGTNCCTAGTTCAAAAAGGNTTGGGCCCAACANATGCC 45

QY 278 CAAGCTGGACTTGGCTTGGGTGCTAAGGTGCAGATGGCCATCT 321

[illegible][illegible]

RESULT 10

BI956651/c

LOCUS	BI956651	816 bp	mRNA	linear	EST 22-OCT
DEFINITION	HVSME0004108f	Hordeum	vulgare	rabchis	EST library HVCDNA0015

(normal) *Hordeum vulgare* cDNA clone HVSME0004J09f, mRNA sequence

ACCESSION	BI956651
VERSION	BI956651 1 GT-16307904

## KEYWORDS

SOURCE	Hordeum vulgare.
ORGANISM	Hordeum vulgare

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae  
: Triticeae. Hordeum

REFERENCE 1 (bases 1 to 816)

**AUTHORS**  
Wing, R., Close, T. J., Kleinhofs, A., Wise, R., Chin, A., Begum, D.,  
Frisch, D., Atkins, M. V., Henry D., Palmer M., Bambo T. S.

J., Oates, R. and Main, D.

**TITLE**  
Development of a genetically and physically anchored EST resource for barley genomics: Morex rachis cDNA library

JOURNAL  
Unpublished (2001)

COMMENT

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA  
Tel.: 864 656 7288

**Fax: 864 656 4293**

Email: [rwing@clemonson.edu](mailto:rwing@clemonson.edu)  
Total bases = 236

Seq primer: AATTAACCCCTCACTAAAGGG

FEATURES  
Location/Qualifiers  
High quality sequence stop: 372.

FEATURED	LOCATION/QUADRANTS
source	1. .816

```
/organism="Hordeum vulgare"  
/cultivar="Morex"
```

```
/cultivar="MOREA"
/db_xref="taxon:4513"
```

```
/clone="HVSME0004J09f"
/clone lib="Hardoum vulcano machis pom 14brave HVSME0004J09f"
```

```
/clone_lib= noideum vulgare iaculis esi libdialy hvcndn
(normal)"
```

```
/tissue_type="Rachis"  
club_height="m1010"
```

177C1C1-135011pap/

/note="Vector: pBluescript SK(-); Site1: EcoRI; Site2: XhoI; Plants were grown at Washington State University, Pullman, WA in a greenhouse, the rachises were excised and frozen in liquid nitrogen (Kleinbolls lab). In the TJ Close lab at the University of California, Riverside total RNA was prepared, poly(A) was purified, one primary unamplified cDNA library was made, and 1 million pfu were in vivo excised to give pBluescript SK(-) cDNA phagemids (Chin). Phagemids were plated and picked at the Clenson University Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see <http://www.genome.clemson.edu/projects/barley>. To order this clone see <http://www.genome.clemson.edu/orders> Also see close TJ, Wing R, Kleinbolls A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (<http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html>)"

BASE COUNT 129 a 257 c 323 g 107 t  
ORIGIN

Query Match 3.6%; Score 180; DB 13; Length 816;  
Best Local Similarity 73.1%; Pred. No. 1.9e-22;  
Matches 247; Conservative 0; Mismatches 85; Indels 6; Gaps 1;

QY 1384 CCCTTGCCAGTCGCCCATCTGTCGCCGCGAGCGGAGGAGCGGATGACA 1443  
Db CCCTTGCCAGTCGCCCATCTGTCGCCGCGAGCGGAGGAGCGGATGACA 1443

QY 1444 G-----GGACATGCGCGAGGAGGAGGCGGTACGAGGAGCATCCGGGACGAG 1497  
Db G-----GGACATGCGCGAGGAGGAGGCGGTACGAGGAGCATCCGGGACGAG 1497

QY 1498 CCGCAGGTGCGGTGAGTCGATCGCTGTGTGAGGCGGCTGTGCGCTCTCCCATCGG 1557  
Db CCGCAGGTGCGGTGAGTCGATCGCTGTGTGAGGCGGCTGTGCGCTCTCCCATCGG 1557

QY 1558 AGCTGCGGAGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1617  
Db AGCTGCGGAGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1617

QY 1618 TCCAGTGGACACCCACACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1677  
Db TCCAGTGGACACCCACACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1677

QY 1678 AACACCACTTGTTCGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1715  
Db AACACCACTTGTTCGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1715

RESULT 11  
BH643146/c 394 bp DNA linear GSS 14-FEB-2002  
LOCUS  
DEFINITION 1008052E03.y1 1008 - RescueMu Grid 1 Zea mays genomic, DNA  
sequence.  
ACCESSION BH643146  
VERSION BH643146.1 GI:18671943  
KEYWORDS GSS.  
SOURCE Zea mays.  
ORGANISM Zea mays

REFERENCE 1 (bases 1 to 394)  
AUTHORS Walbot,V.  
TITLE Maize genomic sequences found using engineered RescueMu transposon  
JOURNAL Unpublished (2001)  
COMMENT Contact: Walbot V

Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Possible ligation site of ends cut by 2 different endonucleases.  
Reverse complemented post-ligation sequence from source sequence.  
Plate: 1008052 row: 18  
Class: transposon-tagged.  
Location/Qualifiers

FEATURES  
source

1..394  
/organism="Zea mays"  
/cultivar="mixed background W23/A188/B73"  
/db\_xref="taxon:4577"  
/clone\_lib="1008 - RescueMu Grid 1"  
/tissue\_type="leaf"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/note="Organ: leaf; Vector: RescueMu (engineered from pBluescript backbone); Site1: BamHI; Site2: BglII; RescueMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site [www.zmldb.iastate.edu](http://www.zmldb.iastate.edu) and follow the links for 'RescueMu.' Grid 1 was grown at Berkeley in 2001. DNA was extracted from leaf punches, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

BASE COUNT 40 a 140 c 159 g 55 t  
ORIGIN

Query Match 3.6%; Score 178.6; DB 17; Length 394;  
Best Local Similarity 74.3%; Pred. No. 3.9e-22;  
Matches 255; Conservative 0; Mismatches 79; Indels 9; Gaps 2;

QY 1371 GTAGAGTGTGCTGCTTGGCCAGTCGTCCTGTCGCCGCGAGCGGCGGAGCGGAC 1430  
Db GTAGAGTGTGCTGCTTGGCCAGTCGTCCTGTCGCCGCGAGCGGCGGAGCGGAC 1430

QY 1431 GAGGCGCATGACAGGAAAC-----TGGCCGAGGAGGAGGAGGAGGAGGAGGAGGAG 1484  
Db GAGGCGCATGACAGGAAAC-----TGGCCGAGGAGGAGGAGGAGGAGGAGGAGGAG 1484

QY 1485 CCGGGGACGACG 1541  
Db CCGGGGACGACG 1541

QY 1542 GGCCTCTCTCCCATCGAGCTGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1601  
Db GGCCTCTCTCCCATCGAGCTGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1601

QY 1602 CGGGGAAACATGCTCCCGTCCAGTGGGACCCACGACGCGCGCGCGCGCGCGCGCGCG 1661  
Db CGGGGAAACATGCTCCCGTCCAGTGGGACCCACGACGCGCGCGCGCGCGCGCGCGCG 1661

QY 1662 CTGCGCTCCCGAGATGAACACACCATCTTTCGACGAGGAGGAG 1704  
Db CTGCGCTCCCGAGATGAACACACCATCTTTCGACGAGGAGGAGGAGGAGGAGGAGGAG 1704

QY 1704 CTGCGCTCCCGAGATGAACACACCATCTTTCGACGAGGAGGAGGAGGAGGAGGAGGAG 1704  
Db CTGCGCTCCCGAGATGAACACACCATCTTTCGACGAGGAGGAGGAGGAGGAGGAGGAG 1704

RESULT 12  
BH443619/c 613 bp mRNA linear EST 25-JUL-2000  
LOCUS  
DEFINITION BH443619.B08.D16ZS Wheat etiolated seedling root normalized cDNA library Triticum aestivum cDNA clone WHE1116\_B08.D16, mRNA sequence.  
ACCESSION BH443619  
VERSION BH443619.1 GI:9443154  
KEYWORDS EST.  
SOURCE bread wheat.



Query Match 3.3%; Score 164.8; DB 17; Length 535;  
Best Local Similarity 85.8%; Pred.No.1.le-19;  
Matches 218; Conservative 0; Mismatches 32; Indels 4; Gaps 3;

QY 4126 TAGAATAGTACTCCCTCGGTCAAAAATAAGTGAGTTTTAGCAGCATTTCAATTGTTCAACG 4185  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 272 TATACTCTACTCCCCTCTGCCAAATAAAGTCAGTTTT-GCACATTTACGGTTCACG 214

QY 4186 TTTAGACAATCGGTCATTATTTGAAGAAAAAGAATAGATAGTAATTTTATTGTTATTAGATG 4245  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 213 TTGTACCCTCGCTCTATTATAAATTTTTTTATGANTAGTATTTTTATTGCTATTAAATG 154  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 4246 ATAAAAACATGAATAGTACATTTTAATGTGCACFAATTTTTTTTAAATATTTTATTAATTT 4305  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 153 ATARAACAATGATGATCTTTATGTGTGACATA--ATAATTTCAAATTTTTCACAAATTT 96

QY 4306 TTCAAATAAGACGGATGGTCAAGCGCTAACATGGAATATCTATGCTACACATTATTTTG 4365  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 95 TTCAAAATAAGACGAGGGTCAAA-CETTTGGACACGGATATCCATGGCTGATCTATTITG 37  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 4366 GGACGGAGGTAGTA 4379  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 36 GGACGGAGGTAGTA 23  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15  
AQ327645/c

LOCUS  
DEFINITION AQ327645 842 bp DNA linear GSS 08-JAN-1999  
nxbx004IB05f CUGI Rice BAC Library Oryza sativa genomic clone  
nxbx004IB05f, DNA sequence.

ACCESSION AQ327645  
VERSION AQ327645.1 GI:4119495  
KEYWORDS GSS.  
SOURCE Oryza sativa.  
ORGANISM Oryza sativa.

REFERENCE  
AUTHORS Wing,R.A. and Dean,R.A.  
TITLE A BAC End Sequencing Framework to Sequence the Rice Genome  
JOURNAL Unpublished (1998)  
COMMENT Contact: Wing RA  
Clemson University Genomics Institute  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu  
Seq primer: TAATACGATCCTCATAGGG  
Class: BAC ends  
High quality sequence start: 58  
High quality sequence stop: 189.

FEATURES  
Location/Qualifiers  
1..842  
    /organism="Oryza sativa"  
    /strain="Japanica"  
    /cultivar="Nipponbare"  
    /db\_xref="taxon:4530"  
    /clone="nxbx004IB05f"  
    /clone\_lib="CUGI Rice BAC Library"  
    /tissue\_type="Leaf"  
    /lab\_host="E. coli DH10B"

Note=Vector: pHELOBACII; Site\_1: HindIII; Site\_2:  
HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of



GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2003, 04:38:30 ; Search time 123.698 Seconds  
(without alignments)  
12398.686 Million cell updates/sec

Title: US-09-702-134-7212\_COPY\_62300\_67300

Perfect score: 5001

Sequence: 1 aggggacactacatttcca.....tgccgcggagtgccgcgcg 5001

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents.NR.\*  
1: /cgn2.6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2.6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2.6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2.6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2.6/ptodata/1/ina/PTUS\_COMB.seq.\*  
6: /cgn2.6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	91.6	1.8	1730	4	US-09-575-574-5
C 2	91.6	1.8	3572	4	US-09-575-574-3
C 3	75.6	1.5	500	4	US-09-575-574-2
C 4	74	1.5	1926	4	US-09-249-585A-4
C 5	74	1.5	1931	2	US-09-130-114-2
C 6	73.2	1.5	1026	4	US-09-105-390-13
C 7	73.2	1.5	2169	4	US-09-105-390-5
C 8	70.6	1.4	7218	1	US-08-332-463-14
C 9	61	1.2	4403765	4	US-09-103-840A-2
C 10	61	1.2	4411529	4	US-09-103-840A-1
C 11	60.6	1.2	1280	4	US-09-060-756-4
C 12	60.6	1.2	13872	4	US-09-105-537-1
C 13	60.4	1.2	1591	1	US-08-728-956-3
C 14	60	1.2	390	4	US-09-197-649-7
C 15	60	1.2	925	3	US-08-858-003-1
C 16	60	1.2	925	3	US-09-078-166-1
C 17	60	1.2	925	4	US-08-997-467-1
C 18	60	1.2	4403765	4	US-09-103-840A-2
C 19	57.6	1.2	1881	4	US-09-434-288-5
C 20	57.2	1.1	4411529	4	US-09-103-840A-1
C 21	56.2	1.1	1590	4	US-09-434-288-1
C 22	56	1.1	1998	6	5212296-8
C 23	54.8	1.1	1138	3	US-08-581-148C-3
C 24	54.8	1.1	1884	1	US-07-704-288C-1
C 25	54.8	1.1	1884	1	US-08-379-259-1
C 26	54.4	1.1	12588	2	US-08-387-942C-1
C 27	54.4	1.1	28958	1	US-08-258-261B-6

C 28	54.4	1.1	28958	1	US-08-456-837-6
C 29	54.4	1.1	28958	1	US-08-457-342-6
C 30	54.4	1.1	28958	1	US-08-457-646A-6
C 31	54.4	1.1	28958	1	US-08-458-076A-6
C 32	54.4	1.1	28958	1	US-08-764-233A-4
C 33	54.4	1.1	28958	1	US-08-457-335A-6
C 34	54.4	1.1	28958	1	US-08-729-214-6
C 35	54.4	1.1	28958	3	US-09-028-934-6
C 36	54.4	1.1	49577	1	US-08-764-233A-1
C 37	54.2	1.1	38506	3	US-09-320-878-19
C 38	54	1.1	6085	4	US-09-029-603-4
C 39	54	1.1	7218	1	US-08-232-463-14
C 40	53.8	1.1	1439	4	US-09-056-556-167
C 41	53.8	1.1	1439	4	US-09-072-596-162
C 42	53.6	1.1	2064	1	US-08-343-428-1
C 43	53.6	1.1	44377	2	US-08-804-227C-7
C 44	53.6	1.1	44377	2	US-08-804-198-1
C 45	53.2	1.1	1140	3	US-09-023-173-4

ALIGNMENTS

RESULT 1  
US-09-575-574-5/C  
; Sequence 5, Application US/09575574  
; Patent No. 6376750  
; GENERAL INFORMATION:

; APPLICANT: Yu, Su-May  
; APPLICANT: Chao, Yu-Chan  
; TITLE OF INVENTION: PLANT SEEDLING AND EMBRYO PROMOTER  
; FILE REFERENCE: 08919-047001  
; CURRENT APPLICATION NUMBER: US/09575,574  
; CURRENT FILING DATE: 2000-05-22  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 1730  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
US-09-575-574-5

Query Match 1.8%; Score 91.6; DB 4; Length 1730;  
Best Local Similarity 65.6%; Pred. No. 2.4e-12;  
Matches 168; Conservative 0; Mismatches 79; Indels 9; Gaps 2;

QY	4171	TATTCATGTTCCACGTTTGACCATTCGCTTATTGAAAAAGATTATGATTAGTATTTT	4230
DB	1352	TTTTCGTCGTCACGTTTAACTGTCGCTTATATGAAAAATTTTATAATTAGTATTTT	1293
QY	4231	TATGTTTATTAGATGATAAAACATGATAGTACTTTATGTGTGACTATTTTATTAAT	4290
DB	1292	CAATGTTGTTTATATGATAAAATATGATTAACTTTATACGTAACCTGCTTTT	1237
QY	4291	ATTTTATTATTTTTCAAATAGACGGATGTCACGCTAAACATGATATCTATG	4350
DB	1236	TTTTTCATTAATTTTTCAAATAATA-----TGATCAAAACATTGGACATGAAACCTCAGG	1182
QY	4351	GCACACATTTTTCGGGCGGAGGTAGTACGCTATAAACACATATAACATATATAA	4410
DB	1181	GTTCGTCCTTTTTCGGGCGGAGGAGTAACTCGTAGTCAGCGGTAGCATACT	1122
QY	4411	GAGATAAAAGAGAAGA 4426	
DB	1121	CAATTATCAATATA 1106	

RESULT 2  
US-09-575-574-3/C  
; Sequence 3, Application US/09575574  
; Patent No. 6376750  
; GENERAL INFORMATION:  
; APPLICANT: Yu, Su-May

```
; APPLICANT: Chao, Yu-Chan
; TITLE OF INVENTION: PLANT SEEDLING AND EMBRYO PROMOTER
; FILE REFERENCE: 08919-047001
; CURRENT APPLICATION NUMBER: US/09/575,574
; CURRENT FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3572
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-575-574-3

Query Match      1.8%; Score 91.6; DB 4; Length 3572;
Best Local Similarity 65.6%; Pred. No. 3.6e-12;
Matches 168; Conservative 0; Mismatches 79; Indels 9; Gaps 2;

QY 4171 TATTCATGTCACGTTTGACCATCGTCTATTGAAAAAGATTATGATTAGTATTTT 4230
      ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1352 TTTTCGTGCCAACGTTTAACTGTCGCTTATATGAAAAATTTTATATAGTATTTT 1293
      ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
QY 4231 TATGTTATTAGATGATAAACAATAGTAGTACTTTTATGTGTGACTAATTTTTTTTAAAT 4290
      ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1292 CATGTTGTTATGATAAATATGATTAATACTTTTATACGTAACCTGGCTTTT----AA 1237
      ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
QY 4291 ATTTTATTATTATTTTCAAAATAGCGGATGTCAAAAGCGCTAAACATGATATCTATG 4350
      ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1236 TTTTTCATAAATTTTTCATATAATA-----TGATCAACAATTTGGACATGAAAACCTCAGG 1182
      ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
QY 4351 GCTACACTTATTTTGGGACGGAGTAGTAGGTATTAACACAGCTATATAACATATTTTAAA 4410
      ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1181 GTTTGCTCTTTTGTGGACGGAGGAGTAAACTCGTAGCTAGCTCAGACGGGTAGCATACT 1122
      ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
QY 4411 GAGATAAAAGAGAGA 4426
      ||||| ||| ||||| |||
DB 1121 CAATTATCAATATA 1106
      ||||| ||| ||||| |||

RESULT 3
US-09-575-574-2/c
; Sequence 2, Application US/09575574
; Patent No. 6376750
; GENERAL INFORMATION:
; APPLICANT: Chao, Yu-Chan
; TITLE OF INVENTION: PLANT SEEDLING AND EMBRYO PROMOTER
; FILE REFERENCE: 08919-047001
; CURRENT APPLICATION NUMBER: US/09/575,574
; CURRENT FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 500
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-575-574-2

Query Match      1.5%; Score 75.6; DB 4; Length 500;
Best Local Similarity 76.2%; Pred. No. 8.7e-09;
Matches 93; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 4171 TATTCATGTCACGTTTGACCATCGTCTATTGAAAAAGATTATGATTAGTATTTT 4230
      ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
DB 122 TTTTCGTGCCAACGTTTAACTGTCGCTTATATGAAAAATTTTATATAGTATTTT 63
      ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
QY 4231 TATGTTATTAGATGATAAACAATAGTAGTACTTTTATGTGTGACTAATTTTTTTTAAAT 4290
      ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
DB 62 CATGTTGTTATGATAAATATGATTAATACTTTTATACGTAACCTGGCTTTTAAATTTT 3
      ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
QY 4291 AT 4292
      |
DB 2 TT 1
```

```
RESULT 4
US-09-249-585A-4/c
; Sequence 4, Application US/09249585A
; Patent No. 6417002
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert
; TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISODES
; FILE REFERENCE: 0867/0D905
; CURRENT APPLICATION NUMBER: US/09/249,585A
; CURRENT FILING DATE: 1999-02-11
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Epstein Barr Virus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1926)
; OTHER INFORMATION: template strand of EBNA-1 DNA
US-09-249-585A-4

Query Match      1.5%; Score 74; DB 4; Length 1926;
Best Local Similarity 47.1%; Pred. No. 4.5e-08;
Matches 227; Conservative 0; Mismatches 255; Indels 0; Gaps 0;

QY 1145 GATPAGAGAAAAAGAGAGAAATATATATAGGGATAATACCTGGATAGAGAGAGTGTG 1204
      ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
DB 980 GGTGGAGACGAGGACGGGAGGACGAGGACGGGAGGAGGACGAGGACGGGAGGACGG 921
      ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1205 GCATGCGGAGATCGCGCGCAGGTGCGGACGAGCATCGCCATCCGCGCGGCGTACTCGT 1264
      ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
DB 920 GGAGGACGAGGACGGGAGGACGGGAGGACGGGAGGACGGGAGGACGGGAGGACGG 861
      ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1265 TGGCGTCGCGCCCGACGCGACGCGTGTGCTCTCTCCCTTGTACAGAGGACAGCGCGATGC 1324
      ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
DB 860 CGGGGAGGACGGGAGGAGGACGAGGACGGGAGGACGGGAGGAGGACGAGGACGGGGA 801
      ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1325 GCGCGCGCGTCTCGAGGCGGACCCCTGGAGCGCGACGAGATCGGAGTAGAGTGGTGC 1384
      ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
DB 800 GGACGGGAGGACGAGGACGGGAGGACGGGAGGACGAGGACGGGAGGACGGGAGGGA 741
      ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1385 CCTTGGCCCGACGCGCCATCCTGTGCGCGACGCGCGGCGGACGAGCGCGATGACAG 1444
      ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
DB 740 CGAGGACGGGAGGACGAGGACGGGAGGACGAGGACGGGAGGACGAGGACGGGAGGA 681
      ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1445 GGAACCTGCGCGGAGCGAAGAGGCGGTAGCGAAGGACATCCCGGGACCGACGCGCGAGG 1504
      ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
DB 680 CGGGGAGGACGGGAGGAGGACGAGGACGGGAGGACGAGGACGGGAGGACGGGAGGA 621
      ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1505 TCGGCTTGGAGTCGATGCCGTTGTGGAGCGGCTCGTGGGCGCTCTCCCTCCATCGGAGTGG 1564
      ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
DB 520 CGGGGAGGACGAGGACGGGAGGAGGACGAGGACGGGAGGACGGGAGGACGGGAGGA 561
      ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1565 GGGAGAGCGGAGGATGAGGGGTTTGGGGCGCGACTCCGGGGGAAACATCGCTCCCAT 1624
      ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
DB 560 GGACGAGGACGGGAGGACGGGAGGAGGACGAGGACGGGAGGACGAGGACGGGAGGA 501
      ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1625 GG 1626
      ||
DB 500 GG 499
      ||

RESULT 5
US-09-130-114-2/c
; Sequence 2, Application US/09130114
; Patent No. 5976807
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert A.
; APPLICANT: Damaj, Bassam B.
; APPLICANT: Robbins, Alan K.
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
```



```

; TITLE OF INVENTION: From Multiple Transfected Episomes
; FILE REFERENCE: 0867/1D903US1
; CURRENT APPLICATION NUMBER: US/09/130,114
; CURRENT FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: EBNA
US-09-130-114-2

Query Match          1.5%; Score 74; DB 2; Length 1931;
Best Local Similarity 47.1%; Pred. No. 4.5e-08;
Matches 227; Conservative 0; Mismatches 255; Indels 0; Gaps 0;

QY 1145 GATAGAGAAAAAGAGAGATAATTAATAGGGGATATATCTGGATGAGGAGAGGTGTG 1204
Db 980 GTGGAGAGACGAGACGGGGAGAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGG 921
QY 1205 GCATGGCGAGATCGGCGCGCAGGTTGCGGACGAGCATGCGCCGCGGCGTACTCGT 1264
Db 920 CGAGGAGAGACGCGGGGAGAGCGGGGAGGACGAGGACGGGGAGGAGGAGGAGGA 861
QY 1265 TGGGTGCGGCCAGCGACGCGTGTGCTCTCCCTCTGTTACGAGGACAGCGCCGATGC 1324
Db 860 CGGGAGAGACGCGGAGGAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA 801
QY 1325 GCGCGCGCGTCTCGAGGCGGACCCCTGAGCGGCGGAGACGAGATCGGAGTAGGTCGTG 1384
Db 800 GGACGGGAGGAGACGAGCGGGGAGGACGCGGGAGGAGGAGGAGGAGGAGGAGGA 741
QY 1385 CTTTGGCCCACTGGCCCTCTCTCTGTCGCGCGGAGCGGCGGAGGAGGAGGAGGAG 1444
Db 740 CGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA 681
QY 1445 GGAATGCGCGGAGGAGGAGGCGGTGAGGAGGAGACATCCCGGAGCGCGCGCAGG 1504
Db 680 CGGGGAGGAGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA 621
QY 1505 TCGGTTGGAGTGATGCGGTTGTCGAGCGGCTGCTGGGCGCTCTCCATCGGAGCTGCG 1564
Db 620 CGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA 561
QY 1565 GGGAGAGCGGAGAGTGGAGGAGGTTGGGCGGCGACTCCGGGGGAGCCATCCGTCACGT 1624
Db 560 GAGCAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA 501
QY 1625 GG 1626
Db 500 GG 499

RESULT 6
US-09-105-390-13
; Sequence 13, Application US/09105390
; Patent No. 6288303
; GENERAL INFORMATION:
; APPLICANT: Rodriguez, Raymond
; TITLE OF INVENTION: Rice Beta-Glucanase Enzymes
; TITLE OF INVENTION: and Genes
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,390
; FILING DATE: Filed herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/050,675
; FILING DATE: 25-JUN-97
; ATTORNEY/AGENT INFORMATION:
; NAME: Petithory, Joanne R.
; REGISTRATION NUMBER: P42,995
; REFERENCE/DOCKET NUMBER: 2000-0455.30

```

```

; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,390
; FILING DATE: Filed herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/050,675
; FILING DATE: 25-JUN-97
; ATTORNEY/AGENT INFORMATION:
; NAME: Petithory, Joanne R.
; REGISTRATION NUMBER: P42,995
; REFERENCE/DOCKET NUMBER: 2000-0455.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; TELEX:
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1026 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-105-390-13

Query Match          1.5%; Score 73.2; DB 4; Length 1026;
Best Local Similarity 75.3%; Pred. No. 5e-08;
Matches 119; Conservative 0; Mismatches 33; Indels 6; Gaps 2;

QY 4375 TAGTAGCTATTAACACGCTATAACATATTTTAAAGAGATAAAGAGAGAGAGAGTAG 4434
Db 283 TATTAGACTATAACACGCTATAACATATTTTAAAGAGATAAAGAGAGAGAGAGTAG 341
QY 4435 CAGCGGACTATATTTTGTAGCCAGCTACAGCAGCGACTCAGAGACGTA-----ATGTGT 4489
Db 342 CAGCGGCTACAGATTGTTAACACCTACAGCAAGACTTTAAGATGATGTGTGTATAA 401
QY 4490 GTATATGACAGTAGGACCGAGGTATTAAAGATAGTAGTA 4527
Db 402 ATCTATGACAGTGGGACCGACGAGCTTATATATATAATA 439

RESULT 7
US-09-105-390-5
; Sequence 5, Application US/09105390
; Patent No. 6288303
; GENERAL INFORMATION:
; APPLICANT: Rodriguez, Raymond
; TITLE OF INVENTION: Rice Beta-Glucanase Enzymes
; TITLE OF INVENTION: and Genes
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,390
; FILING DATE: Filed herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/050,675
; FILING DATE: 25-JUN-97
; ATTORNEY/AGENT INFORMATION:
; NAME: Petithory, Joanne R.
; REGISTRATION NUMBER: P42,995
; REFERENCE/DOCKET NUMBER: 2000-0455.30

```

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2169 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-105-390-5

Query Match 1.5%; Score 73.2; DB 4; Length 2169;
Best Local Similarity 75.3%; Pred. No. 7.6e-08;
Matches 119; Conservative 0; Mismatches 33; Indels 6; Gaps 2;

QY 4375 TAGTAGCTATAACACAGCTATAACATATTTAAAGAGATTAAGAGAGAGAGAGATAG 4434
Db 283 TATTAGACTATAACACAGCTATAACATATTTAAAGAGATTAAGAGAGAGAGAGATAG 4434
QY 4435 CAGCGGCTATATTTTGTAGCAGCTACAGCAGGACTCAAGACGTA-----ATGTGT 4489
Db 342 CAGCGGCTACAGATTGTTAACACCTACAGCAAGACTTTAAGATCATGTGTGTATAA 401
QY 4490 GTATATGACAGTAGGACCAAGGATTAAAGATATAGTA 4527
Db 402 ATCTATGACAGTGGGACGAGCTTAATATATAATA 439

RESULT 8
US-09-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHTEFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION DATE:
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fis
; US-08-232-463-14

Query Match 1.4%; Score 70.6; DB 1; Length 7218;
Best Local Similarity 5.9%; Pred. No. 6.3e-07;
Matches 19; Conservative 194; Mismatches 108; Indels 0; Gaps 0;

QY 1001 GGAAAAAGAGTAGATGTAGATGTAGCAACAACAAGAGAGAGAGAGAGAGAT 1060
Db 1352 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1293
QY 1061 GCCAAGTGCACAGTAAGCTAAACAATCTAGGCGAGCATTCATAGAGGTCAAGCTG 1120
Db 1292 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1233
QY 1121 TATTGATTGGAGAGATATGGATTGATAGAGAGAAAAAGAGAGATATAATAAGGGATA 1180
Db 1232 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1173
QY 1181 ATACCTGGATGAGAGAGAGAGTGTGGCATGGCGATCGCGCAGGTTGCGGACGAGCA 1240
Db 1172 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1113
QY 1241 TGCCCATCGCGGCGGTACTCGTGTGGCGTCGCGCCACGCGCATCGCTCTCCCOCT 1300
Db 1112 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1053
QY 1301 GTACACGAGAGACGCCCGA 1321
Db 1052 CGACCTGCAGCAAGCTCGGA 1032

RESULT 9
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
; US-09-103-840A-2

Query Match 1.2%; Score 61; DB 4; Length 4403765;
Best Local Similarity 47.7%; Pred. No. 0.005;
Matches 244; Conservative 0; Mismatches 260; Indels 7; Gaps 2;

QY 1262 CGTTGGCGTCGGCCACGCGACGCTGTCTCCCTTGTGTACCGAGGAGAGAGAGAG 1321
Db 3928207 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3928148
QY 1322 TGCAGCGCGCGGTCTCTGAGGGCGGACCGCTGGAGGGGAGAGATCGGAGTAGAGTCGG 1381
Db 3928147 CGAATCCGGCGGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGT 3928088
QY 1382 TGCCCTTGGCCCGAGTCGGCGCATCTCTGGTGGCGCGCGCGCGCGCGCGCGCGCGATGA 1441
```

D	b	1191225	C	GTTACCGGGATCAATACCGCTGTTGCCGTTCGCCTTGGCCGCCCGCGCGCGCGCGCGTG	1191284
Q	y	1517	C	GATCCGCTGTGTGGAGCGCTCGTGGGCCCTCTCCCATCGAGCTGCGGGGAGAGCGGA	1576
D	b	1191285	C	CGCGCTCGCGGGTGC	1191344
Q	y	1577	G	ATGAGGGGTGGGGGGCACTCCGGGGAAACAATGCCCTCCACTGGGAGCCCCACCA	1636
D	b	1191345	C	TGTTCGCGCGTTCGCGCGCGCGCGCGCGCGTTCGCGCGCTTCGCGACACC	1191404
Q	y	1637	C	GCCCCCGCGCGCGCATGTTTCGACTTCGCTCCCAGSATGAACACCACTTGTTCGAGC	1696
D	b	1191405	A	GCGCGC-CGACCGCGCGTTCGCACCGCGCGCGCGCGCGCGCTCCGCGCGCACCG	1191463
Q	y	1697	G	AGGAGGGAGACAGCGCGCGCGCGCGCGGAGACAGCAGCAGCAGCAGCATCCGC	1755
D	b	1191464	C	CGACACGCGCTTCGCGAACAATTCGGCGCGTTC	1191522

RESULT 11  
US-09-060-756-4  
; Sequence 4, Application US/09060756

```

1 TITLE OF INVENTION: METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST FROM
2 A SPECIFIC BACTERIAL STRAIN
3 TITLE OF INVENTION: THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED DNA
4 TITLE OF INVENTION: LIBRARY APPLICATION TO THE DETECTION OF MYCOBACTERIA
5 FILE REFERENCE: 3495-0169
6 CURRENT APPLICATION NUMBER: US/09/060,756
7 CURRENT FILING DATE: 1998-04-16
8 NUMBER OF SEQ ID NOS: 743
9 SOFTWARE: PatentIn Ver. 2.0
10 SEQ ID NO 4
11 LENGTH: 1280
12 TYPE: DNA
13 ORGANISM: Mycobacterium tuberculosis
14 US-09-060-756-4

```

[illegible]

; NUMBER OF SEQUENCES: 3

; NUMBER OF SEQUENCES: 3

FILE REFERENCE: NDA02/CL COM  
; CURRENT APPLICATION NUMBER: US/09/197,649

FILE REFERENCE: NDA02/CL COM  
; CURRENT APPLICATION NUMBER: US/09/197,649

```
/ CURRENT FILING DATE: 1998-11-23
/ EARLIER APPLICATION NUMBER: 07/829,461
/ EARLIER FILING DATE: 1992-01-31
/ EARLIER APPLICATION NUMBER: 07/739,055
/ EARLIER FILING DATE: 1991-08-01
/ EARLIER APPLICATION NUMBER: 07/561,968
/ EARLIER FILING DATE: 1990-08-02
/ NUMBER OF SEQ ID NOS: 26
/ SOFTWARE: Patent In Ver. 2.0
/ SEQ ID NO 7
/ LENGTH: 390
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Sequence
/ OTHER INFORMATION: having a 120 repeat of ACG flanked by fixed
/ OTHER INFORMATION: fragments having NcoI restriction sites.
US-09-197-649-7

Query Match      1.2%; Score 60; DB 4; Length 390;
Best Local Similarity 48.0%; Pred. No. 4.4e-05;
Matches 171; Conservative 0; Mismatches 185; Indels 0; Gaps 0;

QY 1183 ACCGTGATGAGGAGAGGTGGCATGCGGATCGCGCGAGCTTCGGACGACATG 1242
Db 31 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 90

QY 1243 GCCATCCGGCGGCGCTACTCTGTCGCGTCCGCCCGACGACGACGACGACG 1302
Db 91 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 150

QY 1303 TACGAGGACGACGCGCGATGCGCGCGCGCGCTTCGAGGCGGACCCCTGG 1362
Db 151 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 210

QY 1363 AGATCGAGTAGAGGTGCGTCCCTTGCGCCAGTCGCGCCATCTCGTCCCG 1422
Db 211 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 270

QY 1423 CAGGGGACGAGCGCGATGACAGGAACTGCGCGGAGGAGGCGGTAGCGAAG 1482
Db 271 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 330

QY 1483 ATCCGGGACGACGCGCGAGTGGGTGGAGTCGATGCTTGTGAGCGGCTC 1538
Db 331 ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 386

RESULT 15
US-08-858-003-1/c
/ Sequence 1, Application US/08858003
/ Patent No. 6060234
/ GENERAL INFORMATION:
/ APPLICANT: Katz, Leonard
/ APPLICANT: Stassi, Diane L.
/ APPLICANT: Summers Jr., Richard G.
/ APPLICANT: Ruan, Xiaolan
/ APPLICANT: Pereda-Lopez, Ana
/ APPLICANT: Kakavas, Stephan J.
/ TITLE OF INVENTION: NOVEL POLYKETIDE DERIVATIVES
/ NUMBER OF SEQUENCES: 34
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Abbott Laboratories
/ STREET: 100 Abbott Park Rd.
/ CITY: Abbott Park
/ STATE: Illinois
/ COUNTRY: USA
/ ZIP: 60064-3500
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
```

```
/ SOFTWARE: FastSeq Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/858,003
/ FILING DATE: 16-MAY-1979
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Dianne Casuto
/ REGISTRATION NUMBER: P-40,943
/ REFERENCE/DOCKET NUMBER: 4952.US.P2
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (847)-938-3137
/ TELEFAX: (847)-938-2623
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 925 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ US-08-858-003-1

Query Match      1.2%; Score 60; DB 3; Length 925;
Best Local Similarity 47.2%; Pred. No. 7.2e-05;
Matches 183; Conservative 0; Mismatches 205; Indels 0; Gaps 0;

QY 1240 ATGGCCATCCGCGGCGCTACTCTGTCGCGTCCGCCCGACGACGACGACG 1299
Db 707 ATGCGCTGGTGGAGGTGGTGGAGGTGTCTGCTAGCGACGCGGTGCGTATGC 648

QY 1300 TGTATACAGAGACGACGCGCGATCGCGCCCGCGTCTCGAGGCGACCCCTG 1359
Db 647 TGGTCCCACTGGTGTGTGCGGGGTGGGTGCGCGGTGAGCATCTGATCAGGG 588

QY 1360 ACGAGATCGGAGTAGAGTTCGCTGCTTGGCCCGATCTCGTCCCGCGACG 1419
Db 587 GTGTGCGGCGGTGGTGGGTGAGGTGTCTGAGGTGGTGGTGGTGGTGGTGG 528

QY 1420 GCGCAGGGGACGAGGCGCGATGACAGGGAACCTGCGCGGACCGAAGAGG 1479
Db 527 TGGTGGCGATGTGTGAGGAGGTGTGCGGTGCGCGGAGAGACGCGTGGCG 468

QY 1480 GACATCCCGGACCGACCGCGAGTGGGTGGAGTCCGATCCCGTTGTGAG 1539
Db 467 TTGATGCGCGCATGGAGATGGTGTCTGTGGTGTCTGAGGTGGGTGGGTGG 408

QY 1540 TGGGCTCTCTCCCATCGGAGTGGCGGAGAGGCGGAGGATGGAGGGTTGG 1599
Db 407 GGTGCGTGGCGACGCGGACGATGCGCGCGGCGGCGAGGTTTTCATGAGG 348

QY 1600 TCCGGGGGAACCATGCCGTCCCACTGGG 1627
Db 347 CGTGTGTGACGAGGTGGACGCGCGTCGG 320

Search completed: March 26, 2003, 18:50:29
Job time : 20400.7 secs
```

Result	No.	Score	Query		Length	DB	ID	Description
			Match					
	1	121.8	2.4	5757	10	US-09-993-170-56	Sequence 56, Appl	
	2	118.4	2.4	949	10	US-09-770-445-342	Sequence 342, App	
C	3	114.4	2.3	547	9	US-10-101-4634-374	Sequence 374, App	
C	4	111.2	2.2	2000	10	US-09-887-576-870	Sequence 870, App	
	5	83.2	1.7	2000	10	US-09-887-576-859	Sequence 859, App	
	6	68.2	1.4	520	9	US-10-184-644-332	Sequence 332, App	
	7	61	1.2	863	10	US-09-974-300-5249	Sequence 5249, Ap	
C	8	60.6	1.2	15872	9	US-09-860-846-1	Sequence 1, Appli	
	9	60.6	1.2	15872	10	US-09-861-289-1	Sequence 1, Appli	
	10	60	1.2	390	10	US-09-790-399-7	Sequence 7, Appli	
	11	60	1.2	925	9	US-09-735-050-1	Sequence 1, Appli	
C	12	58	1.2	615	9	US-10-125-815-4	Sequence 4, Appli	
C	13	57.4	1.1	2010	12	US-10-032-717-9	Sequence 9, Appli	
C	14	56.4	1.1	405	10	US-09-878-574-3891	Sequence 3891, Ap	
C	15	55.2	1.1	2087	10	US-09-349-385-11	Sequence 11, Appl	
	16	54	1.1	284	10	US-09-369-373-472	Sequence 472, App	
	17	54	1.1	2000	10	US-09-887-576-832	Sequence 832, App	
C	18	53.2	1.1	4689	9	US-09-860-846-34	Sequence 34, Appl	
C	19	53.2	1.1	4689	10	US-09-861-289-34	Sequence 34, Appl	

RESULT 2  
US-09-770-445-342

; Sequence 342, Application US/09770445  
; Patent No. US20020023281A1  
; GENERAL INFORMATION:  
; APPLICANT: Goriach, Jörn  
; APPLICANT: An, Yong-Qiang  
; APPLICANT: Hamilton, Carol M.  
; APPLICANT: Price, Jennifer L.  
; APPLICANT: Raines, Tracy M.  
; APPLICANT: Yu, Yang  
; APPLICANT: Rameake, Joshua G.  
; APPLICANT: Page, Amy  
; APPLICANT: Matthew, Abraham V.  
; APPLICANT: Ledford, Brooke L.  
; APPLICANT: Woessner, Jeffrey P.  
; APPLICANT: Haas, William David  
; APPLICANT: Garcia, Carlos A.  
; APPLICANT: Krieker, Maja  
; APPLICANT: Slader, Ted  
; APPLICANT: Davis, Keith R.  
; APPLICANT: Allen, Keith  
; APPLICANT: Hoffman, Neil  
; APPLICANT: Hurban, Patrick  
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis  
; FILE OF INVENTION: thaliana  
; FILE REFERENCE: 2023US (PARA-012PRV)  
; CURRENT APPLICATION NUMBER: US/09/770,445  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: US 60/178,472  
; PRIOR FILING DATE: 2000-01-27  
; NUMBER OF SEQ ID NOS: 999  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 342  
; LENGTH: 949  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-770-445-342

Query Match 2.4%; Score 118.4; DB 10; Length 949;  
Best Local Similarity 55.8%; Pred. No. 5.9e-20;  
Matches 249; Conservative 0; Mismatches 191; Indels 6; Gaps 1;  
QY 1183 ACCTGATGAGGAGAGGTGTGGCATGGCGGAGATCGGCGGAGTGGCGGAGGATGGCGGAGGACATG 1242  
DB 400 ACCTGAATGATGGGAAGAGAGGTGTGATGATCATGACGAGGATCTTAATCAAAAGA 459  
QY 1243 GCCATCGCGGGCGGTACTCGTTGGCGTGGCCGACGCGAGGTGTGGCTCTCCCTGG 1302  
DB 460 TCCATATTGTTCCGTTAGCTCTCGGCTCATGATGATCCACACGTCACTCTCTCTTGA 519  
QY 1303 TACCAGAGACAGCCGATGCGCGCGGCTCGAGGGGACCCCTGGAGCGCGGAGC 1362  
DB 520 TACCACACACCGCTTGTATCTTCGCGGCAATTCCTACTCTCTCTCTCTCTTGACC 579  
QY 1363 AGATCGAGTAGAGTTCGGTCCCTTGGCCAGTGGCCATCTGTGGCGCGGAGGCG 1422  
DB 580 ATCTCTCTGACAGTGTCTCCAGCTCCACTCTTTTTCGCGGTTCACCGGAGCG 639  
QY 1423 CAGGGACGAGGCGGATGACAGGAACTGGCGGAGCGAAG-----GAGGGGTAGG 1476  
DB 640 CAGGGACCAACCGGATCACAGCGAATCTGTTTCCAGCGGATCTTCAACCGGTTAGG 699  
QY 1477 AAGGACATCCGAGACCGACCGCAGGTGCGTTGGAGTCGATGCGGTTGTGGAGCGC 1536  
DB 700 AACGCCATCTCTGACCTACCTCCACACACTTACCTGTGTCAATGTCAACGTGTAGTGC 759  
QY 1537 TCGTGGGCTCTCTCCATCAGGTGCGGGGAGAGCGGAGGATGAGGGGTGGGGGCG 1596  
DB 760 TCGTGTGCTCTCTCCACCGGAGATCTGGAGCGGAGGATGATGATGATGATGATG 819  
QY 1597 CACTCCGGGGAGAACCATGCGCTCCA 1622  
DB 820 CATTCCGGTGAAGGATTTATCCA 845

RESULT 3  
US-101-101-464A-374/C  
; Sequence 374, Application US/10101464A  
; Publication No. US20030046728A1  
; GENERAL INFORMATION:  
; APPLICANT: Strabala, Timothy  
; APPLICANT: Nieuwenhuizen, Nicolaas  
; APPLICANT: Higgins, Colleen M.  
; TITLE OF INVENTION: Compositions Isolated from Plant Cells  
; FILE OF INVENTION: and Their use in the Modification of Plant Cell Signaling  
; FILE REFERENCE: 11000.1020G2  
; CURRENT APPLICATION NUMBER: US/10/101.464A  
; CURRENT FILING DATE: 2002-03-18  
; PRIOR APPLICATION NUMBER: 09/704,302  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: 09/228,986  
; PRIOR FILING DATE: 1999-01-12  
; PRIOR APPLICATION NUMBER: 60/162,866  
; PRIOR FILING DATE: 1999-11-01  
; PRIOR APPLICATION NUMBER: PCT/US00/00724  
; PRIOR FILING DATE: 2000-01-11  
; NUMBER OF SEQ ID NOS: 989  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 374  
; LENGTH: 547  
; TYPE: DNA  
; ORGANISM: Pinus radiata  
US-101-101-464A-374

Query Match 2.3%; Score 114.4; DB 9; Length 547;  
Best Local Similarity 52.9%; Pred. No. 4.2e-19;  
Matches 275; Conservative 0; Mismatches 236; Indels 9; Gaps 1;  
QY 1183 ACCTGATGAGGAGAGGTGTGGCATGGCGGAGATCGGCGGAGTGGCGGAGGACATG 1242  
DB 546 ACCTGATGAGGAGAGGATCGGATGTTTCAGATCAGAGCGGATATCATGATAGTCTC 487  
QY 1243 GCCATCGCGGGCGGTACTCGTTGGGTGGCCGACGCGAGCTGTGGCTCTCCCTGG 1302  
DB 486 TCGAGATATGTTATAGCTATAGCATTTCTTCAGATTTCTGTATCACTTTCGCTTGA 427  
QY 1303 TACCAGAGACAGCCGATGCGCGCGGCTCTGAGGGGACCCCTGGAGCGGCGGAGC 1362  
DB 426 TACCAGAGAAATGCTCTTTAGAACGCTCCATGATTCAGAGCTGCTTGTCTTAGTATC 367  
QY 1363 AGATCGGAGTAGAGTCTGCTTGGCTTGGCCAGTGGCCATCTGTGGCGCGGAGGCG 1422  
DB 366 ATATTCCTTAGAGTGTCTCTCTTTCCCATTTCTTTCATCGCGGTACCTCCATGCA 307  
QY 1423 CAGGGACGAGCGGATGACAGGAACTGGCGGAGCGA-----AGGAGGCGTTA 1473  
DB 306 CAGGGAACATAAACCAATACGAGCTTTTGGGATCTTTATATGTTTCAGTATAGTGT 247  
QY 1474 GCCAAGGACATCCGGGACGAGCGCGAGTGGGTTGGAGTCGATGCGGTGTGGAGC 1533  
DB 246 GCAATATCATTCAGAGGACTTCTCCACACACTTGTGTATCTATATCGGATGCAAA 187  
QY 1534 GGCTCTGGGCTCTCTCCATCGAGCTCGGGGAGAGCGGAGAGTGGAGGTTGGG 1593  
DB 186 GGAACATGAGCTGTCTCCCAATGAAGATGTAGTTGAGCTAAGAAATGAAGGTGTGT 127  
QY 1594 GCGACTCCGGGGAGAACATGCGTCCAGTGGGACCCACCGCCCGCCCGCCGCC 1633  
DB 126 TTGCATCTTTGAGGAAGAAATGATCCCAATGAGAACCATGAATCTCTCCCTGGCCGAC 67  
QY 1654 ATGTTGACTGCTCCCGAGGATGAACACACACTGTTTCG 1693  
DB 66 ATGTTGCTCTGCTTCGAGAGAGAGATATCCATTTTCG 27

RESULT 4  
US-09-887-576-870/C

```

RESULT 5
US-09-887-576-859
  Sequence 859, Application US/09887576
  Patent No. US20020144047A1
  GENERAL INFORMATION:
  APPLICANT: Budworth, P.
  APPLICANT: Brown, D.
  APPLICANT: Chang, H.
  APPLICANT: Zhu, T.
  APPLICANT: Han, B.
  APPLICANT: Wang, X.
  APPLICANT: Cooper, Bret
  TITLE OF INVENTION: Promoters for re
  FILE REFERENCE: 1360.001US1
  CURRENT APPLICATION NUMBER: US/09/88
  CURRENT FILING DATE: 2001-06-25
  PRIOR APPLICATION NUMBER: US 60/213,
  PRIOR FILING DATE: 2000-06-23
  PRIOR APPLICATION NUMBER: US 60/214,
  PRIOR FILING DATE: 2000-06-23
  PRIOR APPLICATION NUMBER: US/06/258,
  PRIOR FILING DATE: 2000-12-29
  NUMBER OF SEQ IDS NOS: 875
  SOFTWARE: FastSeq for Windows Version

```

Query Match      1.4%; Score 68.2; DB 9; Length 520;  
Best Local Similarity    20.1%; Pred. No. 3.2e-07;  
Matches     89; Conservative    91; Mismatches    259; Gaps        0;

QY   1086   AATCTAGGCGAGCATTCATAGAAGTCCAGCGTGATTGGATTGGAGAGATATGCATTG   1145  
         : : : : : | : : : : : | : : : : :  
Db   1   MNKKIKEDLSSTQQAFHQIAMEPFINVPKRRNGVNFLAVVVIIYLITAGA   60

QY   1146   ATAGCGAATAAAAGCAATAATAATAGGCAATAACCTGGATCAGCGAAGGTGTGG   1205  
         : : : : : | : : : : : | : : : : :  
Db   61   GLVVQVINLQARLVLEMYFLNDTLAADSPSLLQSAPHEHLAQASRLQLVAOL   120

QY   1206   CATGCCGAGATCGGCGCCAGCGATTGCGGACAGCATGCCATCCGGCGGGCGTACTCGTT   1265  
         : : : : : | : : : : : | : : : : :  
Db   121   TWWRVSHHLLORVDNFTONGPMFRIGKQGQAQGLIOGHKGAMGFAPGPDPGPPAEKGA   180

QY   1266   GSCTCGSCCTCAGCGCACGCTGTGCTCTCTCCCCTGTGCTACTACAGAACAGCGCTCGTAGCG   1325



```

Db 181 GAWGRDAGATGSPGQGPVGVKAGLQGPQAPRQKQATGTPQGEKSKDGLIGPK 240
QY 1326 GCGCGCGGTCTCGAGGCGGACCTCGAGCGGACGAGATCGAGTAGAGTCTCGGTGCC 1385
Db 241 GTGTGKGEKDLGLPSKSGDRGMKGDAGVMGPFGAQSCKDFGRPPGLAGFPKAGDQ 300
QY 1386 CTTGGCCACGTCGCGCATCTCTGTGTCGCGCCGACGCGGACGAGGACGCGCATGACAGG 1445
Db 301 GQFGLQGVGPPGAVGPHFAGKFPFGAGSPGRAGLPSPGSPGATGLKSKGDTGLQOQ 360
QY 1446 GAACGTGGCGGAGGAGGAGGCGCTTAGCGAAGGACATCCGCGACCGACGCCGAGGT 1505
Db 361 GRKSGVPGVAGVGTGGQSGFLAGPKGAPQAGKQDQGVKSSGEGQGVKGEKGERGEN 420
QY 1506 GCGGTTGGAGTCGATGCC 1523
Db 421 SVSVRVGSSNRGAEEV 438

RESULT 7
US-09-974-300-5249
; Sequence 5249, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE OF INVENTION: Expression
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5249
; LENGTH: 863
; TYPE: DNA
; ORGANISM: Bacillus clausii
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(863)
; OTHER INFORMATION: n = A,T,C or G
US-09-974-300-5249

Query Match 1.2%; Score 61; DB 10; Length 863;
Best Local Similarity 49.0%; Pred. No. 3.3e-05;
Matches 216; Conservative 0; Mismatches 219; Indels 6; Gaps 2;

QY 1208 TGGCGAGATCGCGGCGAGGTTCGCGACGAGATGCGCATCCGCGGCGGTACTCGTTGG 1267
Db 213 TGACGCGACGCGCATGCCACGCGGATCGGACGCCGACGCGCATGCCGATGCCGAGC 272
QY 1268 CPTGCGCCAGCGACGCTGTGCTCTCCCTGTTACGAGACAGCGCGCATCGGCG 1327
Db 273 GGACGCTGTATCCGACGCTGATGTCACGC---GGATCCGACGCGGACGCGATCGGA 329
QY 1328 CGCGGCTCTTCGAGGCGACCCCTGAGCGGCGGACGAGATCGGAGTAGAGTCTCGTGCCT 1387
Db 330 TGGCGACGCCGATGCGACGCTGACGCGGATGCGGATGCGGACCGGATGCGGACGCGA 389
QY 1388 TGGCCACGTCGCGCATCTCTGTGCGCGCGGCGGACGAGGACGCGCGCATGACAGGA 1447
Db 390 TGCACGCTGACGCTGACGCTGATGCGCGATGCGGATGCGGATGCGGATGCGGACGCGA 449
QY 1448 ACTGCGCGGAGGAGGAGGCGCTAGCGAGGACATCCGCGGACGCGACGCCGAGGTGC 1507
Db 450 TGCNGATGCGGATTCGGACGCCGATGCCGATGCCGATGCCGATGCCGATGCCGATGCC 509
QY 1508 GGTGTGAGTCGATGCGGTTGTGAGAGCGGCTGCTGGGCGCTCTCTGCC---ATCGGAGTCGG 1564

```

```

Db 510 TGTGACCGGATGCCCATGCCGATCGGATCGGACGCCGATCGGATCGGATCGGATCGGA 569
QY 1565 GGGAGAGCGGAGGATGGAGGCTTGGGGCGGCACCTCCGGGGGAACCATGCGCTCCCACT 1624
Db 570 CGCTATCGGAGCCCATCGGATGCTGACGCGACGCCGATCCGATCCGATCGACGCTGA 629
QY 1625 GGGACCCCAACACGCCGCCGC 1645
Db 630 CGCTGATGCCATGCNGACGC 650

RESULT 8
US-09-860-846-1/c
; Sequence 1, Application US/09860846
; Patent No. US2002016472A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/860,846
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 15872
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-860-846-1

Query Match 1.2%; Score 60.6; DB 9; Length 15872;
Best Local Similarity 45.5%; Pred. No. 0.00039;
Matches 216; Conservative 0; Mismatches 259; Indels 0; Gaps 0;

QY 1207 ATGGCGAGATCGGCGCGAGTTGGGAGCAGCATGGCCATCCGCGGCGGTACTCGTTG 1266
Db 5238 ACGACCACGCGGTGGTTCGTTGAGCGGCGCGGACCTTSCCTCTGAGCCCTTC 5179
QY 1267 GCGTCGGCCAGCAGCGCTGCTCTCCCTGTTACAGAGGACAGCGCGGATGCGG 1326
Db 5178 AGGACTCGAGACATCGGCTCGCCGCTGACGCGGATCATCGCGCGGATCGCGCG 5119
QY 1327 CGCCCGGTCTCGAGGCGACCCCTGGAGCGCGGACGAGATCGGAGTAGAGTCTCGTCC 1386
Db 5118 GCGGACTGATGAGCGCGCGCGGCGGTGACGAGACGTGGCGGTCTCGAGGAGAGG 5059
QY 1387 TTGGCCCACTCGSCCATCTTGTGTCGCGCGGCGGAGGAGGCGGATGACAGG 1446
Db 5058 ACACGGGACGTGCGCGGCGGATCTCGCGGACGAGATGGCGGCGGTGAGCAGGTGCG 4999
QY 1447 AACTGCCCGGAGCAAGGAGGCGGTAGGAAAGGACATCCCGGAGACCGCGCAGGTG 1506
Db 4998 ACAGGCGCTGTGTTCGAGGCGGAGACGCGCCACTCGAAGGCGAAGAGGCGGCG 4939
QY 1507 CGGTTGAGATCGATCGCGCTGTGTGAGCGGCTCTGTTGGGCTCTCTCCCATCGAGTCTCG 1566
Db 4938 TGGGTGTACGCGTGTGAGTCTGAGTCTCTCCCGGCGGAGATCTCCCGACGCGCG 4879
QY 1567 GAGAGCGAGATGAGGAGGTTGGGGGCGCACTCCGCGGAGACCATGCGTCCGATGCG 1626
Db 4878 TCGAGTGTGAGTGTGAGGCGCGGTACACGCGTGTGAGGCGGCGGCGGAGACAGG 4819
QY 1627 GACCCACACGCCGCCCGCGCGCATGTTTCGACTGCGCTCCCGAGGATGAACA 1681
Db 4818 GCGGCGCGAGTCTCTGCCCCATGGGACGCGCTGGGCTCTCTGCGCGGTGAACA 4764

RESULT 9

```

```

US-09-861-289-1/c
; Sequence 1, Application US/09861289
; Patent No. US20020110897A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/861,289
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 15872
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-861-289-1

Query Match      1.2%; Score 60.6; DB 10; Length 15872;
Best Local Similarity 45.5%; Pred. No. 0.00039;
Matches 216; Conservative 0; Mismatches 259; Indels 0; Gaps 0;

QY 1207 ATGGCGAGATCGCGCGGAGGTTCGGGACGAGCATGGCCATCCGGGGGCGTACTCGTTG 1266
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5238 ACGACACGACGGGTGGTCCGTGTGACGGCGGACGGGACCCCTCCCTCGTAGCCCTTC 5179
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1267 GCGTCGGCCAGCGGCGACGTGTCGCTCCCTCGGTACGAGGACACGCCCGCATCGG 1326
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5178 AGGAGCTCGACGACCTCGGCTCGCCCGCTGCAAGCGCATCATCGCCCGCCCTCGCG 5119
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1327 CGCGCGGTCTCGAGGGCGACCTCGAGGGCGGACGAGATCGGAGTAGAGTTCGGTCCC 1386
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5118 GCCAGCTGCATCGACGGCGCGCGGCGGTGACGAGACGTGCGCGCTCGTCGAGGAGAG 5059
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1387 TTGGCCAGTCGCGCATCTGCTGTCGCGCGGCGGCGGACGAGCGCATCACAGGG 1446
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5058 ACACGGCGACGTGCGCGCGCGGCTGTCGCGGAGGTGCGCGGTGAGCAGTTCGGGG 4999
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1447 AACTGGCGGAGGAGGAGGCGTTCGAGGAGGACATCCCGGACCGACGCCCGAGGTG 1506
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4998 ACGAGGCGGTGTGTTGAGGAGGCGGAAACAGCGCCACCTCGAAGCGAGGCGGCG 4939
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1507 CGGTGAGTCGATCGCGCTGTTGAGGCGGCTCTGCGGCTCTCCCATCGAGTCCGG 1566
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4938 TGGGTGTACGCGGTGAGGTCCAGCTCCCTCCCGGCGGCGACGATCTCCCGACGGCG 4879
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1567 GAGAGCGGAGGATGAGGGGTGTTGGGCGCGCATCTCCGGGGGACCATCGCTCCAGTGG 1626
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4878 TCGAGGTGACGCTCGAGGCGCGGTACACGCTGTCGAGGCGCGCGGCGAACAACGGGTGC 4819
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1627 GACCCGACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1681
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4818 GCGGCGCGCATCTCCCGCCCATCGCGAGCGGCTGCGCTCCCTGCGCGGTGAACA 4764
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
US-09-790-399-7
; Sequence 7, Application US/09790399
; Patent No. US20020038000A1
; GENERAL INFORMATION:
; APPLICANT: Gold, Larry
; APPLICANT: Turk, Craig
; APPLICANT: Smith, Jonathan D.
; TITLE OF INVENTION: Systematic Polypeptide Evolution by Reverse Translation
; FILE REFERENCE: NEX02/CI-CON2
; CURRENT APPLICATION NUMBER: US/09/790,399
; CURRENT FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: 09/197,649

```

```

; PRIOR FILING DATE: 1998-11-23
; PRIOR APPLICATION NUMBER: 07/829,461
; PRIOR FILING DATE: 1992-01-31
; PRIOR APPLICATION NUMBER: 07/739,055
; PRIOR FILING DATE: 1991-08-01
; PRIOR APPLICATION NUMBER: 07/561,968
; PRIOR FILING DATE: 1990-08-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 390
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Sequence
; OTHER INFORMATION: having a 120 repeat of ACG flanked by fixed
; OTHER INFORMATION: fragments having NcoI restriction sites.
US-09-790-399-7

Query Match      1.2%; Score 60; DB 10; Length 390;
Best Local Similarity 48.0%; Pred. No. 3.3e-05;
Matches 171; Conservative 0; Mismatches 185; Indels 0; Gaps 0;

QY 1183 ACTGTGATGAGGAGAGGTGTGGCATGGCGAGATCGCGCGAGGTTCGGGACGAGCATG 1242
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31  ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 90
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1243 GCCATCGGGGGGGTACTTCGTTGGCGTTCGGCCCGACGCGACGCGTGTGCTCTCCCTCG 1302
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 91  ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 150
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1303 TACGAGGAGGACGCGCCGATCGGCGCGCGTCTCGAGGGGACCGCTGGAGCGCGGAG 1362
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151  ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 210
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1363 AGATCGGATAGAGTTCGCTTCCTTCGCGCCAGTCGCGCATCTGTCGCGCGCGAGGG 1422
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 211  ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 270
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1423 CAGGGGACGAGCGCCATGACAGGGAACGTGCGCGGACGAGGAGGGCGTTAGCGAGGAC 1482
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 271  ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 330
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1483 ATCCCGGACGACGCGCGAGTGGTGGTGGATCGATCCGCTTGTGAGCGGCTC 1538
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 331  ACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 386
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
US-09-735-056-1/c
; Sequence 1, Application US/09735056
; Publication No. US20030013662A1
; GENERAL INFORMATION:
; APPLICANT: Katz, Leonard
; APPLICANT: Stassi, Diane L.
; APPLICANT: Summers Jr., Richard G.
; APPLICANT: Ruan, Xiaolan
; APPLICANT: Pereda-Lopez, Ana
; APPLICANT: Kakavas, Stephan J.
; TITLE OF INVENTION: NOVEL POLYPEPTIDE DERIVATIVES
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Rd.
; CITY: Abbott Park
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS

```

```

SOFTWARE: FastSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/735,056
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/858,003
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dianne Casuto
REGISTRATION NUMBER: P-40,943
REFERENCE/DOCKET NUMBER: 4952.US.P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (847)-938-3137
TELEFAX: (847)-938-3623

```

Query Match	1.2%;	Score 60;	DB 9;	Length 925;
Best Local Similarity	47.2%;	Pred. No. 6.4e-05;		
Matches 183;	Conservative	0; Mismatches 205;	Indels 0;	Gaps 0;
QY 1240	ATFGCCATCCGGGGCGGTACTCGTTGGCTGCGCGCCACGCGACGGTGTGCTCTCCCCC	1299		
Db 707	ATGCCGTGGTGGAGGTGGTGGAGGGTCTCGGTCTAGCGGACGGGTCCGGTCAATGTC	648		
QY 1300	TGGTACACAGAGACACGCCGATACGGGCCCGCGGTCTCGAGGGCGACCTCGAGCGCGGG	1359		
Db 647	TGGGTCCAGTGGGTGGTTCGGGGTGGGTGGGTGGCGGTGAGCATGCTGATAGGGGT	588		
QY 1360	ACGAGATCGGAGTAGAGGTTCGGTGCCTTGGCCACGTCGCGCCATCCTGGTCCCGCCGACG	1419		
Db 587	GTGTGTGGGGGTGGTGGTGGAGGGTGTGAGGGTGGTGGTGAAGGGTTGGTTGAGT	528		
QY 1420	GCACGAGGGACGAGGCCGATGACAGGGAACATGGCCGAGCAAGGAGGGCGCTTAGCGAAG	1479		
Db 527	TGGGTGGCATGTGTTGGAGGGTGGTTCGGGTCTGCCGAGGAGGACGGTGGCTGGGGCCG	468		
QY 1480	GACATCCGGGACCGACCGCGAGGTGGGTGGAGTCGATGCCGTGTGGAGCGGGCTCG	1539		
Db 467	TTGATGGCGGCATGGAGATGGTTCGTGGTGGTGTTCGAGGTGGGGTGGAGGGTGTGG	408		
QY 1540	TGEGGCTCTCCCATCGAGTTCGGGGAGAGGCGGAGGATGGAGGGTTGGGGGCGCAC	1599		
Db 407	GGGTGGTGGGACCGCGAGCATGGCGCCCGGGGGCGAGTTTTCATAGGGGTCCG	348		
QY 1600	TCCGGGGAAACCATTCGCTCCCGATGGG	1627		
Db 347	CGTGTGGTGCACGAGTGGACGGCGTCGG	320		

```

RESULT 12
US-10-125-815-4/c
; Sequence 4, Application US/10125815
; Patent No. US20020173008A1
; GENERAL INFORMATION:
; APPLICANT: Hu, Zhihao
; APPLICANT: McDaniel, Robert
; APPLICANT: Santi, Daniel V.
; TITLE OF INVENTION: RECOMBINANT POLYKETIDE SYNTHASE GENES
; FILE REFERENCE: 300622004720
; CURRENT APPLICATION NUMBER: US/10/125,815
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 09/679,279
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: 60/190,024
; PRIOR FILING DATE: 2000-03-17

```

```

; PRIOR APPLICATION NUMBER: 60/158,305
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: megX56
US-10-125-815-4

Query Match      1.2%; Score 58; DB 9; Length 615;
Best Local Similarity 48.2%; Pred. No. 0.00015;
Matches 163; Conservative 0; Mismatches 175; Indels 0; Gaps

QY 1316 CGCCGATGGCGGCGCGGTCTCAGAGGGCGACCTTGAGGGCGGAGACGAGATCGAGTGA 1375
      ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 498 GCGCGCGCGCGGCGAGCGCGTGCACGCGCGACCGCGCGAGCGCGGACCGCTGCGCGACGA 439
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1376 GTGCGGTGCCTTGCGCCAGTGTGCGGCATCTCTGTGCGCGCGCGAGCGGAGGACGACG 1435
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 438 CCAGGCGCGGACGACACCACTGTCAGCAGCGCGCTGAGCGCGCGCGGACAGACCATCGGCG 379
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1436 CGATGACAGGGAACCTGGCGGAGCGAAGAGAGGCGCTTAGCGAAGAGCAATCCGCGGACCGA 1495
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 378 CGACACCGCGCGCGCGAGCGCGAGACGACCTTGATACCCCGACGACGCGCGCGCGCG 319
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1496 CGCCGCGAGGTGCGGTGTGAGTTCGATCCCGTTGTGAGCGGCTGTGGGCGCTCTCCCATC 1555
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 318 CCTGAGCGTGGCGGAGCTGTGCGCTTGACCTGCCGACGACCGCGGCGCGAGCGCGCGCG 259
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1556 GGAGCTCGGGGAGAGCGGAGGATGGAGGGGTGGGGGCGCGACTCCCGGGGGAACCATGC 1615
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 258 GCGCGACCGCGTACTGTCCGAGGAGGCGCCACGCTGACGGGATCCCGGAGCGGCTGC 199
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1616 GGTCCCGAGTGGAGCCCCACACGCGCCCCCGCGCGCCCGCGC 1653
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 198 CGGTGCGGTGGGCTCGACGACTCCGAGCTCGCGCGCC 161
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 13
US-10-032-717-9/c
; Sequence 9, Application US/10032717
; Patent No. US20020151709A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Nicholas B. Duck
; APPLICANT: Xiang Feng
; APPLICANT: Ronald D. Flanagan
; APPLICANT: Theodore W. Kahn
; APPLICANT: Lynn B. Sims
; TITLE OF INVENTION: Genes Encoding No. US20020151709A1el Proteins With
; TITLE OF INVENTION: Pesticidal Activity Against Coleopterans
; FILE REFERENCE: 35718/237005
; CURRENT APPLICATION NUMBER: US/10/032.717
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 60/242,838
; PRIOR FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2010
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2010)
; OTHER INFORMATION: Manze optimized Cry12I8-1
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: mol2181-1
US-10-032-717-9

```

Query Match 1.1%; Score 57.4; DB 12; Length 2010;  
Best Local Similarity 44.5%; Pred. No. 0.00054;  
Matches 229; Conservative 0; Mismatches 286; Indels 0; Gaps 0;

QY 1208 TGCGAGATGCGCGCAGGTGCGGACGAGCATGCCATCCGCGGCGTACTCTGTTGG 1267  
DB 862 TGGGAGAGAGGCCACACCGTCGAGCACGCGGAGGCTCATCTCCGCGCGAACTGGTGT 803  
QY 1268 CGTCGCGCCAGCGCAGGTGTGCTCTCCCTCTGTTACGAGGACGACGCGCATCGCGC 1327  
DB 802 AGTCCACCCACTGCTTGGCGGAGGTGCCCTTGAGCTTGGCGAGGCGGTCTCATACCACT 743  
QY 1328 CGCGGCTCTCGAGGCGGACCTCGAGCGCGGACGAGATCGGATAGAGTCGGTCCCT 1387  
DB 742 TCAGCGAGTGGTTCGGAGTACTCGCGGCTGAGCTTCATCTGCGCTGCTGATAGTTGTTGA 683  
QY 1388 TGCGCCAGTCGCGCCATCTGTCGCGCGCACGCGCGAGGACGAGCCGATACAGGGA 1447  
DB 682 TGTGTGCTGTCGACGACCCGCTCTCCGCGGAGGAGGCGTCTTGGAGGAGGA 623  
QY 1448 ACTGCGCGGAGGAGGAGGCGCTTAGCGAAGACATCCCGGGACCGACCGCGAGTGC 1507  
DB 622 GGTGGAGGTTGCGGCGCCATGCGCTTACACGCTGAGGAGGCGACCTCGAAGTTGTTCAACG 563  
QY 1508 GGTGTGAGTCGATGCGCTGTGAGCGGCTGCTGGGCGCTCTCCATCGGAGCTCGGGG 1567  
DB 562 GGAAGGAGGCGATGACTGCTGAGGAGGAGTTCGAGGATCGAAGCGGTTGCGCACT 503  
QY 1568 AGAGCGGAGGATGAGGAGGTTGGGCGCGCACTCCGCGGGGAAACATCCCTCCAGTGG 1627  
DB 502 CGCGGAGGCGCGGAGCGCTTGGGCTTCTCTCCACTCTCCAGGCGGTGAGGTAGA 443  
QY 1628 ACCCGACGAGCGCGCGCGCGCGCGCTGCTGAGTGCCTCCCGAGGAGAACACCACT 1687  
DB 442 GCTGTGAGTGTGTCGCGAGGCGCTTCCAGCTCGGAGAGGCGCTTGTTCGCGGCGTACTCGG 383  
QY 1688 TGTTCACGAGGAGGAGGACACGCGCGCGCGCC 1722  
DB 382 CGAICTCTGCTGTGATGAGCTCTCCACTGCTCC 348

RESULT 14  
US-09-878-574-3891/c  
; Sequence 3891, Application US/09878574  
; Patent No. US20020110548A1  
; GENERAL INFORMATION:  
; APPLICANT: Byrum, Joseph R.  
; APPLICANT: Thompson, Michael D.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(15401)B  
; CURRENT APPLICATION NUMBER: US/09/878,574  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 09/333,535  
; PRIOR FILING DATE: 1999-06-14  
; NUMBER OF SEQ ID NOS: 15775  
; SEQ ID NO 3891  
; LENGTH: 405  
; TYPE: DNA  
; ORGANISM: Glycine max  
; OTHER INFORMATION: Clone ID: LIB3028-004-Q1-B1-E11  
US-09-878-574-3891

Query Match 1.1%; Score 56.4; DB 10; Length 405;  
Best Local Similarity 47.5%; Pred. No. 0.00029;  
Matches 166; Conservative 0; Mismatches 186; Indels 0; Gaps 0;

QY 1243 GCATCCGCGGCGTACTGCTGTCGCGCGCAGCGAGTGTGCTCTCCCTCGG 1302  
DB 405 GCTAGCACGAGGAGAAACCGCGGAGGCGGTAGGAACCTTGACGGATTTCGCGCGG 346

QY 1303 TACCAGAGACAGCGCGCATCGCGCGCGGCTCTCGAGGCGACCCCTGGAGCGCGCGAGC 1362  
DB 345 TTGCGATCGAGGCGGAGGTGAGGAGATGATGCGGATGCCACCAAGGACAAACATGAAT 286  
QY 1363 AGATCGGATAGAGTGTGCTGCTTGGCCCATGTCGCGCATCTCTGTCGCGCGACCGCG 1422  
DB 285 CCGAGGAGAGGCGCTCGATGCTACTGCGGCTTGACGCGCGGACATGAGAGAGCG 226  
QY 1423 CAGGGAGCAGGCGCATGACAGAGAACTGCGCGGAGCGAAGAGGCGGTTAGCGAAGAC 1482  
DB 225 GCGCGAGCGCGCGCTGTAGGGTCTCTGAGTGGAGCGCATGCGGGGCGCTCGACGATG 166  
QY 1483 ATCCCGGAGCGACCGCGCGAGTTCGGTGGAGTGCATGCGGTTGTGGAGCGCTCTG 1542  
DB 165 ATCTGTAACAATGCGGAGACGACCATGAAGTAGGTGAGGAGGAGGAGGCGAAGAGC 106  
QY 1543 GCCTCTCTCCATCGGAGTTCGCGGAGCGGAGGATGAGGGGTTGGGGCG 1596  
DB 105 GCGTTCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 52

RESULT 15  
US-09-349-385-11/c  
; Sequence 11, Application US/09349385  
; Patent No. US20020152495A1  
; GENERAL INFORMATION:  
; APPLICANT: Ito, Toshiro  
; APPLICANT: Fromm, Michael  
; APPLICANT: Meyerowitz, Elliot  
; TITLE OF INVENTION: PLANTS HAVING SEEDLESS FRUIT  
; FILE REFERENCE: MBI-0002  
; CURRENT APPLICATION NUMBER: US/09/349,385  
; CURRENT FILING DATE: 1999-07-09  
; EARLIER APPLICATION NUMBER: 60/115,967  
; PRIOR FILING DATE: 1999-01-15  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 2087  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: CDNA  
US-09-349-385-11

Query Match 1.1%; Score 55.2; DB 10; Length 2087;  
Best Local Similarity 44.9%; Pred. No. 0.002;  
Matches 210; Conservative 0; Mismatches 258; Indels 0; Gaps 0;

QY 1194 GAGAAGGTGTGCGATGCGGAGATCGCGCGAGGTTGCGGAGGAGCATGCCATCCGCGG 1253  
DB 1069 GAGATGCGGACCATGCTGCTGCTCGCGAGCTCTCTGCACCCCTCGAGGAGGAGGAGC 1010  
QY 1254 GCGGTACTGCTTGGCGTGGCGCCAGCGACGCTGCTCTCCCTCTGCTGCTGCTGCTG 1313  
DB 1009 GTCGAGGAGTCAAGCTGCTGCTGAGGCGCGGAGTTTGGCGGCGCGCGCTGCTC 950  
QY 1314 AGCGCGGATGCGCGCGCTCTCGAGGCGGCGCTGAGGCGGCGGAGGAGGAGTCA 1373  
DB 949 GTCGATGAGCGCACGAGGAGGAGTCTGGAGCGCGCGGACGCGCGCGCGCGGCT 890  
QY 1374 GAGGTGCGTTCCTTGGCGCATCTGCTGCTGCGCGGAGCGCGGAGGAGGAGGAGGAG 1433  
DB 889 GACGTTGCTTGGTCTGCTGAGTTCGCGGAGCGGAGGAGGAGTCTGCGACCATGTAAGGC 830  
QY 1434 GCGATGACAGGAGAACGTCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1493  
DB 829 GCGGAGGAGTCAAGCGCTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 770  
QY 1494 GAGCGCGAGGTCGCGGTCGAGTTCGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1553  
DB 769 GCGCGCTGATGTGAGCTGTACCGCGTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 710

QY 1554 TCGAGCTCCGGGAGAGCGGAGGATGGAGGGTTGGGGCGCACTCCGGGGGACCAT 1613  
 Db 709 GCGGCGCTGGAGGTTCGGGCGGAGGCGGACGGCGCCAGAGCCGACTGTTCGGCGGCGAT 650  
 QY 1614 GCGTCCAGTGGGACCCACACGCGCCCGCGCCCGCCCATGTTCGA 1661  
 Db 649 GAGCGGAGCATGGCCTCCGCTCACCTTGGGTCCTCCCGGCTCGTGGGA 602

Search completed: March 26, 2003, 14:11:21  
 Job time : 412.743 secs

US 097522920MP1



Creation date: 14-08-2003  
Indexing Officer: PRT1634 - GAU1634 PRINTER  
Team: 1600PrintWorkingFolder  
Dossier: 09752292

Legal Date: 12-08-2003

No.	Doccode	Number of pages
1	CTNF	9
2	DIST	1
3	SRFW	1

Total number of pages: 11

Remarks:

Order of re-scan issued on .....